

平成16年度バイオ人材育成事業  
(アノテーター人材)

報 告 書

平成17年1月

委託先 バイオインフォデザイン・ジャパン株式会社

国際バイオインフォマティクス研究所

委託元 三井情報開発株式会社

## 目次

---

1 . 事業の概要 .....	1
1 . 1 事業体制 .....	2
1 . 2 事業内容 .....	3
2 . 育成対象とする人材像 .....	5
2 . 1 育成対象とする人材のイメージ .....	5
2 . 1 . 1 アノテーターの人材像 .....	5
2 . 1 . 2 アノテーターの業務形態 .....	6
2 . 1 . 3 キャリアパス .....	7
2 . 2 当該人材が必要とされる背景 .....	9
3 . 事業の経過 .....	11
3 . 1 スキルスタンダードの作成 .....	11
3 . 1 . 1 作成手順 .....	11
3 . 1 . 2 作成過程 .....	11
3 . 2 人材ニーズ調査 .....	15
3 . 2 . 1 調査の概要 .....	15
3 . 2 . 2 調査結果 .....	15
3 . 3 カリキュラムの作成 .....	19
3 . 3 . 1 概要 .....	19
3 . 3 . 2 スキルスタンダードからカリキュラムの科目への生成過程 .....	19
3 . 3 . 3 レベル設定 .....	21
3 . 3 . 4 カリキュラムからのシラバスの作成 .....	22
3 . 3 . 5 必要科目の検討 .....	22
3 . 3 . 6 カリキュラムの実用性及び汎用性 .....	23
3 . 3 . 7 次年度以降のカリキュラム・シラバスの活用 .....	23
3 . 4 実証プログラムの実施 .....	24
3 . 4 . 1 実証プログラムの概要 .....	24
4 . スキルスタンダード .....	31
4 . 1 スキル項目とスキル・レベル .....	31
4 . 1 . 1 育成対象とする人材像 .....	31
4 . 1 . 2 スキルスタンダードの基本構成 .....	31

4 . 2	スキルスタンダード	33
4 . 2 . 1	配列解析 (シーケンス)	33
4 . 2 . 2	発現プロファイル解析 (DNA マイクロアレイ・DNA チップ)	36
4 . 2 . 3	多型解析 (SNP)	40
4 . 2 . 4	プロテオーム解析 (質量分析)	43
4 . 2 . 5	タンパク質間ネットワーク (遺伝子ネットワーク) 解析	46
4 . 2 . 6	資質・志向	49
5 .	カリキュラム	51
5 . 1	教育コース	51
5 . 2	カリキュラム	52
5 . 2 . 1	アノテーション・スタンダード (アノテーター初級) 人材育成コース	52
5 . 2 . 2	アノテーション・エキスパート (アノテーター中級) 人材育成コース	54
5 . 2 . 3	アノテーション・マネージャー (アノテーター上級) 人材育成コース	57
5 . 3	シラバス	60
6 .	スキルスタンダード・カリキュラムの活用について	87
6 . 1	スキルスタンダード・カリキュラムの活用方法	87
6 . 2	次年度以降の展開方針	90

#### 参考資料

A.	バイオ産業を支える人材の育成に関する調査 (アノテーター) 結果	91
B.	アノテーター実証講座講義記録	103
C.	アノテーター育成実証講座終了テスト	117
D.	実証講座モニター受講者毎回アンケート	130
E.	実証講座モニター受講者事前アンケート	139
F.	実証講座モニター受講者事後アンケート	140

# 第1章 事業の概要

---

本事業は、バイオとITの融合領域であるバイオインフォマティクス(生命情報科学)を担う具体的な人材像としてアノテーション人材(アノテーター)の育成システムを開発したものである。

## 【アノテーション人材(アノテーター)概説】

アノテーション(annotation)という言葉を英和辞典で引くと、「注釈」、「注釈を付けること」とある。バイオインフォマティクスにおいては、生物(生命)情報に注釈を付けることと定義することができる。その例として、DNA配列情報に生物学的な注釈を付加することが挙げられる。DNA配列は普通4種の塩基(A, C, G, T)の羅列で表現されるが、この文字の羅列を見ただけではその生物学的な意味を汲み取ることはできない。そこで、最初から最後までそれぞれの塩基に番号を付けて、何番から何番までの塩基は特定の生物学的意味をもっているなどと注釈を加えるのである。生物学的意味には多くの種類があって、それも研究が進む度に増加したり、修正されたり、或いは削除されたりする。その中には、遺伝子の始めや終わりを表すものもあれば、その遺伝子の働きを示すものもある。このように、位置と意味を注釈として加えることにより、情報としての価値が始めて出てくるといっても過言ではなく、また第三者にも利用できるようになる。

DNA配列は非常に多くの塩基で構成されているので、とても人間が目で見えて調べるわけにはいかない。そこでコンピューターに配列情報や注釈情報を貯蔵しておいて、コンピューターに解析させたり、解釈させたりすることになる。このような情報をコンピューター処理に適するようにして蓄えておき、コンピューターネットワーク上で公開すれば、世界の誰もが利用できるようになる。このような情報の貯蔵をデータベースといい、DNA配列情報のデータベースとしては、日・米・欧の3カ所で共同構築している国際塩基配列データベースが世界的に知られている。日本では、国立遺伝学研究所の日本DNAデータバンク(DDBJ)が担当している。

アノテーションはDNA配列に限らず、タンパク質や遺伝子発現など多くの生物情報についても行われている。その共通するところは、生物学的な注釈を付加することにある。注釈の質、量そして新しさがそのデータの価値を決定することになる。

具体的なアノテーションの作業は、既知遺伝子情報とのホモロジー(配列類似性)解析、遺伝子産物であるタンパク質のモチーフ解析、遺伝子の属性分類等の方法が用いられる。世界標準になりうるような高品質のアノテーションを付与するためには、遺伝学、免疫学、がん、神経系等の専門家が、提示された結果を検証する作業が不可欠となり、アノテーション人材のトップクラスもここに位置するが、実務的には、あるレベル以上の生物学または生命科学専攻の人材が、コンピューターを駆使することで可能なものでもある。

## 1 1 . 事業体制

本事業は、経済産業省が三井情報開発（株）への委託事業として実施している「平成16年度バイオ人材育成事業」の再委託事業の1つとして実施しているものである。

本事業を実施するにあたり、事業内容の定期的なチェックと作成スキルスタンダード、カリキュラム、シラバスに対する専門的かつ第3者の視点を反映するために、アノテーション分野及びバイオインフォマティクス分野における著名な学術有識者、製薬企業の研究責任者からなる有識者委員会を組織し、非常に有効で活発なご議論をいただいた。この場をお借りして、有識者委員会のメンバーの先生方には、心より感謝申し上げたい。

以下に、有識者委員会のメンバーの諸先生方と具体的に実務を進めた実務担当者のメンバーを記載する。

### 【有識者委員会】

- ・ 委員長  
五條堀 孝（国立遺伝学研究所教授・生命情報・DDBJ研究センター長）
  
- ・ 委員  
館野 義男（国立遺伝学研究所教授）  
石和 貞男（国際バイオインフォマティクス研究所社長、  
お茶の水女子大学名誉教授）  
田中 博（東京医科歯科大学教授）  
菅原 秀明（総合研究大学院大学教授）  
水島 洋（国立がんセンター研究所室長）  
不破 亨（湧永製薬株式会社 副社長）

### 【実施担当者】（中心業務を記載）

- ・ 原案作成、検証&フォローアップ、実証関係、他  
水野 政彦（バイオインフォデザイン・ジャパン株式会社）  
小見山 智義（株式会社国際バイオインフォマティクス研究所）  
花岡 秀樹（株式会社国際バイオインフォマティクス研究所）
  
- ・ 企画調整、連絡調整、報告関係、他  
塩見 均（バイオインフォデザイン・ジャパン株式会社）  
保科 哲也（バイオインフォデザイン・ジャパン株式会社）  
月岡 敬二（株式会社国際バイオインフォマティクス研究所）

## 1 2 . 事業内容

本アノテーション人材（アノテーター）育成事業は、以下のフローで進められた。各項目の詳細内容は、後述する。

- 1 ) 「アノテーター・スキルスタンダード」の作成
  - ・ 実務担当者による原案作成
  - ・ 現職アノテーターへのヒアリング
  - ・ 有識者委員会での討議、修正
- 2 ) 「人材ニーズ調査」の実施とスキルスタンダードへの反映
  - ・ 三井情報開発株式会社と共同でスキルスタンダード項目に関して人材ニーズ調査を実施。
  - ・ 結果分析と、スキルスタンダードへ反映
- 3 ) 「アノテーション技能開発カリキュラム」の作成
  - ・ 実務担当者による原案作成
  - ・ 有識者委員会での討議、修正
  - ・ シラバスの作成
- 4 ) 「実証講座・演習（カリキュラム抜粋版）」の実施
  - ・ 現職アノテーター、若手研究者による演習
  - ・ アンケートの実施とスキルスタンダード・カリキュラムへのフィードバック
- 5 ) 「確認テスト」の実施
  - ・ 実証講座内容に基づいた確認テストの実施
  - ・ 結果分析とアンケート結果の分析及び反映
- 6 ) 有識者委員会による結果総括と次年度事業の方向性検討
  - ・ 実務担当者による作成と報告
  - ・ 報告書及び最終経費報告書の作成



## 第2章 育成対象とする人材像

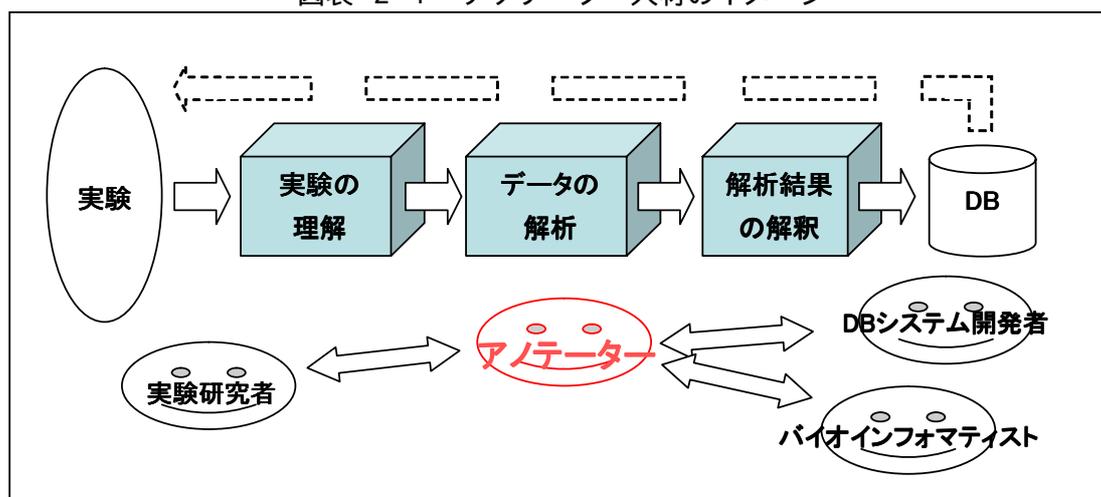
### 2.1 育成対象とする人材のイメージ

#### 2.1.1 アノテーターの人材像

ポストゲノム時代を迎え、大量の生物データが生み出されている。その後の研究および応用利用においてこれらを有効活用するために、大量データを処理できるコンピューターを用いた情報科学が不可欠となり、生物学と情報科学を融合させた研究領域であるバイオインフォマティクスの重要性が認識されるようになった。このバイオインフォマティクス研究はデータベース等のデータをもとに結果を導き出すことから、データに付与されている説明部分 - アノテーション - が極めて重要になる。また、実験から得られる一次データはそのままでは利用することが困難であり、医学・薬学・農学を含む生物学上の機能の特定や意味付け、すなわちアノテーションをすることが非常に重要になってきた。こうしたアノテーションは生物データの分析指標となり、バイオインフォマティクス研究のみならずバイオ研究の礎となる。

アノテーターは、このような重要なアノテーションを行う人材 - 「バイオ研究・バイオ産業の基礎となる生物データの意味付け、分類整理、データベース作成をおこなう人材」である。アノテーターは大量の生物データが生み出されるところに不可欠である。大量のデータに対して、正確かつ一貫したアノテーションを行うためには、生データの信頼性も評価できる実験自体の理解、検索・予測等解析ツールの知識・活用といったデータの解析力、そして分子生物学と情報科学に支えられた解析結果の解釈力という高度な能力が必要である。

図表 2-1 アノテーター人材のイメージ



人材育成という観点では、「実験の理解」、「データ解析」、「解析結果の解釈」の順に教育することが望まれる。これら3つの能力、「実験の理解」、「データ解析」、「解析結果の

解釈」はアノテーションの流れの中で、単純な一方通行ではなく、フィードバックが行われ互いに影響し合う密接な関係にある。研究プロジェクトの中では、アノテーターはこれら3つの能力を駆使して、実験研究者、バイオインフォマティクス研究者、およびデータベースシステム開発者らと円滑にコミュニケーションし、連携し、研究プロジェクトを推進する原動力となることが期待される。

## 2 1 2 アノテーターの業務形態

### (1) 専任アノテーターの職場

前述のように、アノテーターは大量の生物データが生み出されるところに、今後必須であることから、現在活躍している、そして今後活躍すると想定される職場は以下のとおりである。

- ゲノムセンター、シーケンス機関（大学、企業）
- X線結晶解析・NMRセンター
- DDBJ、KEGG等DBの入力・作成者

本事業では、このような高度な専任のアノテーターの育成を最終目標として、これを上級レベルのアノテーターと位置づけた。専任のアノテーターには、データベースの国際的な基準への準拠が強く求められることから、国内外の研究機関との連携、最新文献の理解などの国際性も必須である。また、国際的なデータベースに依拠した優れた解析能力による解析結果を、研究者にフィードバックすることで研究指針に関与することにより、双方向性を推進できるチームでの協調性やコミュニケーション力、先端の科学に対する進取性（前進性）そして個人情報等データの機密・保護に要求される高い倫理性など、資質及び志向面で求められるものも多い。

### (2) 兼任アノテーターの職場

専任のアノテーターとしてその作業を追及する専門職は、今後のバイオ産業の進展とともに必須であるものの、上級であるが故にその必要数は限られている。しかしながら、アノテーションを必要とする実験データの流れを再度検証してみると、アノテーションの能力は実験がおこなわれ、一定以上のデータが産み出され、それがデータベース化されるのであれば、必要な能力であることがわかる。また、現在、多くの研究所・大手バイオ関連企業においても自社データベースの構築が進められていることが予想される。本事業をアノテーション技能のスキルスタンダード開発として捕らえれば、すなわち「アノテーション能力」を考えれば、より幅広い層への適用が可能となる。本事業では、専任アノテーターのレベルを上級とする一方で、企業内で生物実験もこなしながらアノテーション能力を駆使し生物データベースの構築に関与する層を兼任アノテーターとし、初級及び初級から中級へのレベルとして想定した。

さらに、上記の流れの中で、データベースへの格納、構築部分を除外すると、実験から

の大量のデータを受けて、実験内容の理解 データの解析 解析結果の解釈となり、これは、バイオテクノロジーの実験をある一定以上の規模でおこなっているほとんどの研究室にとって必須の技術であるとも捕らえうる。これはバイオインフォマティクスの一つの能力と重なる部分でもあるが、逆に多くの生物系・生命科学系の教育を主に受けた人材から見れば、アノテーション能力の習得からバイオインフォマティクスの世界を臨む方が自然の流れであり、はるかに習得しやすい道のりであるといえる。したがって、アノテーション能力の入門から中級レベルの段階をバイオインフォマティクス習得のファースト・ステップとして捉えることもできる。例えば、アノテーションの能力をある程度身につければ、どのような状態の細胞からどのような遺伝子・タンパク質を得たのかなどの実験によって生データが生み出された状況・背景を把握している者がそのままデータ解析、データの解釈を行うことも可能となる。

このようなアノテーション能力を習得した人材が活躍できる職場の例として、以下のようものが考えられる。

- 製薬会社から DB 開発を受注しているソフトウェア開発会社
- SNP タイピングのデータ入力者（大学、バイオベンチャー、CRO（臨床試験受託機関））
- 質量分析器の測定・データ処理者
- DNA チップのデータ処理者
- 質量分析器、DNA チップのメーカー、販売会社
- CRO での血液サンプルの診断データの測定・処理者
- 細胞を用いた化合物スクリーニングのデータの測定・処理者（製薬会社）
- 教育者（高校教師、専門学校講師、大学バイオ学科講師）
- IT 系コンサルティングファーム（製薬、化学、食品担当）
- 組替え植物の生育・収穫データの DB 化（植物バイオ企業）
- ダイエット食品のモニター試験を請け負う CRO
- 特許に関するデータベース検索が必要な製薬会社の知財管理室（物質特許）

そこで本事業では、初級を終了し中級に進むレベルとして上記の層を捉え、スキルスタンダードの初級、中級レベルの人材像として想定した。

## 2 1 3 キャリアパス

### (1) 専任アノテーター

本事業は、アノテーター育成システム開発事業として、高度な専任のアノテーターを育成するために、初級から上級までのスキルスタンダードを設定し、専任のアノテーターを目指す人が、初級レベルから段階を経て上級レベルへと到達できるよう設定した。

### (2) 兼任アノテーター

専任アノテーターの育成が重要であるのと同様に、既に特定の専門分野を持つ人材がアノテーション能力を身につけることの重要性を認識し、その際の目安となるようにレベルを設定した。これはアノテーターを必要とする場合は多岐にわたっているからであり、例え

ば、治験にかかわる CRO、医師、食品メーカーのバイオ実験研究者等の人材がそれぞれの専門・バックグラウンドを生かしながら、データも解析、注釈をつけることができることが今後ますます必要になると想定されるからである。

加えて重要なことは、バイオインフォマティクスとの関わりである。バイオインフォマティクスは、日本ではその重要性は認識されているものの事業への適用が遅れている。また、バイオインフォマティクスはカバーする分野は幅広く、新規参入や初学者にはそれを始めるきっかけをつかむことが難しく、その具体的なイメージもつかみにくい。数式が出てくるなど大学院以上での専門として学ばなければ、実用に耐える能力になりにくい。しかし、アノテーター人材ならば、具体的な実験から入って、それぞれの分野またはカリキュラムのコースに沿って学ぶことができる。また、対象となる実験の種類が明確なため、生物実験者が具体的なイメージを持って参加しやすいだけでなく、必要な実技演習も絞り込んで実施しやすい。したがって、生物学・生命学からのバイオインフォマティクス初学者にとって、バイオインフォマティクスを初めて学ぶ最初の導入に適切である。つまり、今各人が持っている専門の知識や経験を最大限に生かしながら、バイオインフォマティクス分野へシフトする最も現実的かつ簡便な入り口がアノテーション能力であるといえるだろう。アノテーション能力を習得した人材がアノテーション領域に留まらず、あるものはバイオインフォマティクス研究者へ、あるものは IT 技術者との架け橋となって、バイオと IT の融合を促進し、バイオ研究・バイオ産業が活性化する一助となることが大いに期待される。

## 2 2 . 当該人材が必要とされる背景

ポストゲノム時代を迎え、遺伝子やタンパク質のゲノムワイドな解析実験から大量の生物データが生み出されている。これら一次データはそのままでは利用することが困難であり、医学・薬学・農学を含む生物学上の機能の特定や意味付け、すなわちアノテーションをすることが非常に重要になってきた。こうしたアノテーションは生物データの分析指標となり、バイオインフォマティクス研究のみならずバイオ研究の基礎となる。正しいアノテーションが行われなければバイオインフォマティクス研究はできないと言っても過言ではないだろう。

真に有効なアノテーションを行うために、アノテーターには特に以下の3つの能力が必要である。

- 1) 生データの信頼性も評価できる実験自体の理解力
- 2) 解析プログラムの知識・活用といったデータの解析力
- 3) 主に分子生物学と情報科学に支えられた解析結果の解釈力

また、アノテーターはこれまでの研究者にはない状況に置かれている。大量データを統一した基準で処理する必要があるため、与えられるデータ、解析手順、データベースへの入力項目等、業務の枠組みがおおよそ決まっており、自由に実験、解析する研究者の状況とは異なる。一方、解析方法の選択や解析結果の解釈に必要な検索データベースの知識や各解析方法の特徴の把握等のスキルがアノテーターに特に求められている。上級レベルにおいては、アノテーターは実験研究者、バイオインフォマティクス研究者等、他の研究者やデータベースシステム開発者と共同して業務を進める必要から特定の専門分野に限らず幅広い知識とコミュニケーション能力を必要とし、アノテーターがこれらの構成員の中で橋渡し役として中心的な役割を果たすことが期待される。そのような実験とバイオインフォマティクスを結ぶ包括的な視座は進展著しいバイオ研究のリーダーに求められるものであり、医療・産業への応用も含めてアノテーターはポストゲノム時代になくしてはならない存在となるであろう。

バイオインフォマティクスを理解する人材の緊急性が言われてから久しく、国内の大学や専門学校ではバイオインフォマティクスおよびアノテーターを育成するカリキュラムは欧米に比べて極めて少ないと言える。ゲノム解析以降、配列解析というまでもなく、DNAチップによる発現プロファイル解析、質量分析によるプロテオーム解析等、大量の実験データが生み出されてきた。今後、そのデータ生産のペースはさらに加速することが予想される。一方で、それらの生データに注釈を加え、分類整理し、データベースの形で保存する人材、すなわちアノテーター人材は実験室の現場でほとんど見受けられない。大量のデータが蓄積していく過程で、データをどのように有効活用するのか、またデータベースからバイオインフォマティクスのアプローチで、データに込められた生物学的意味を発掘または再構築するのかがわが国のバイオ研究・バイオ産業にとって緊急の課題である。



## 第3章 事業の経過

---

### 3 1 . スキルスタンダードの作成

#### 3 1 1 作成手順

スキルスタンダードは以下の手順で作成した。

- 1) 作成チームでの第1案の検討
  - ・人材像を設定し、業務領域や業務内容を検討
- 2) 専任アノテーターへのヒアリング(2004年6月)
  - ・ヒアリングを反映して、スキルスタンダード基本骨子の策定
- 3) 知識項目の洗い出し
  - ・骨子に沿って、具体的な知識項目を列挙
- 4) 第1回有識者委員会の実施(2004年7月)
  - ・スキルスタンダード基本骨子、知識項目・スキル項目の検討
  - ・レベル設定の検討
  - ・有識者委員会での知見を知識項目・スキル項目へ反映
- 5) 人材ニーズ調査の実施(2004年8月)
  - ・知識項目・スキル項目についての意見集約
  - ・レベル設定についての意見集約
- 6) 知識項目・スキル項目の完成
- 7) 中間報告会
  - ・報告会での要請事項の検討と反映
- 8) 第2回有識者委員会の実施(2004年10月)
  - ・カリキュラム・シラバスを含む意見の検討と反映
- 9) 実証講座の実施
  - ・実証講座モニター結果及びアンケート等の知見の反映
  - ・カリキュラム・シラバスへの知見の反映
- 10) 外部ヒアリングの実施
  - ・ほぼ完成のスキルスタンダードに対する意見
  - ・次年度事業展開への助言
- 11) 第3回有識者委員会の実施(2005年1月)
  - ・完成スキルスタンダードの承認
- 12) 最終報告会の実施(2005年2月)
  - ・完成スキルスタンダードの報告

#### 3 1 2 作成過程

##### (1) 実験分野を柱とする業務領域の分析

スキルスタンダードの素案を作成するにあたり、アノテーターは大量の生物データが生み出されるところに必要であるという人材像を仮定した。スキルスタンダードまたはカリキュラムの作成の際、例えば統計学という科目が考えられるが、その内容は広範にわたり、初学者が学ぶのにどこから始めればよいかわからないだけでなく、個々の項目が実際の研究・業務にどう直結するのかはわかりにくい。そこで、導入口としてより具体的な各実験分野に着目し、スキルスタンダードのスキル項目・知識項目の最も基本的な分類を実験分野別とする構成を試みた。実験分野の選定にあたってはデータ処理にコンピューターが不可欠となるような大量のデータを生み出し、ポストゲノム研究の中で重要かつ代表的な実

験分野を以下のように列挙した。

【実験分野】

ゲノムシーケンス、cDNA シーケンス、X 線結晶解析、NMR、発現分布・発現解析（DNA チップ・マイクロアレイ）、SNP タイピング、質量分析、タンパク質間相互作用（two hybrid system、プロテインチップ）

このうち、ニーズが高いと予想され、かつ大量のアノテーションが課せられるハイスループットな実験に着目し、業務内容として選ばれたものは以下の 5 つの実験分野である。

【選択された 5 つの実験分野】

- 1) 配列解析（シーケンス）
- 2) 発現プロファイル解析（DNA マイクロアレイ・DNA チップ）
- 3) 多型解析（SNP）
- 4) プロテオーム解析（質量分析）
- 5) タンパク質間ネットワーク（遺伝子ネットワーク）解析

また、文献データベースからのテキスト・マイニング等、データベースに登録済みのデータを基にアノテーションを行なう場合も検討されたが、実験室から生み出される生データを扱うという観点から、本事業では対象の業務領域としては実験分野に重点を置いた。この第 1 案の 5 つの実験分野はヒアリング、有識者委員会で不可欠だと承認された。また、これらの実験分野に加え、「資質・志向」（国際性等）の重要性が指摘され、6 番目の柱としてスキルスタンダードに反映された。

（2）データの流れた業務フローの分解

業務領域の大分類として選択した実験分野を分析した結果、いずれの実験分野も共通して、実験装置からデータベースへの入力までのデータの流れて、「実験内容の理解」、「データの解析」、「解析結果の解釈」の 3 段階に分けることができると判断した。「実験内容の理解」は「データの解析」、「解析結果の解釈」の前提となるものであり、業務領域・業務内容の分類の一つのため、便宜的に業務フローに含めた。これら業務フローの 3 段階の必要性について、実験分野ごとに人材ニーズ調査で確認し、さらにヒアリングで検証、修正を行った。

（3）スキル・レベルの設定

人材像から想定される職種・専門の育成対象者から最低 3 段階のレベルが必要と考え、素案として初級・中級・上級の 3 段階に分けた。有識者委員会およびヒアリングではそれぞれがどのようなことができ、どう異なるのかというレベル設定を検討した。有識者委員会では特に上級レベルの専任アノテーターについての育成に焦点が当てられ、初級から中

級を経て上級レベルへステップアップを図るキャリアパスが提案された。また、アナテーター自身が結果の利用または応用という出口について認識している必要性が指摘された。上級レベルはプロジェクトリーダーであり、中級レベル以下を指導するというリーダーでもあるという意見も踏まえ、結果の解釈において中級では「総合判断」、上級では解析結果からの応用利用をさらに加え「総合判断・産業応用」のスキルを設定した。ただし、初級ではこの項目は割愛し、重点を実験の理解やデータ解析に置くこととした。これらのレベル設定は人材ニーズ調査のアンケートで9割以上の支持を得た。

#### (4) スキルスタンダードの基本構造の決定

以上より、スキルスタンダードの基本構造を、5つの実験分野および「資質・志向」を柱とし、各実験分野は実験内容の理解、データの解析、解析結果の解釈の3段階に分け、かつスキル・レベルとして初級・中級・上級の3段階に分類したものとした。

#### (5) スキル項目・知識項目の特定

知識項目は、各実験分野とも共通して、業務内容中の業務フローの3段階に対応して以下のように細分化された。

- 1) 実験内容の理解：「実験目的、実験原理の知識」、「実験装置、実験方法の知識」
- 2) データ解析：「解析ツールの知識(検索データベースの知識)」、「解析ツール(検索・予測ツール)の原理の知識」、「解析ツールの利用法の知識」
- 3) 解析結果の解釈：「分子生物学の知識」、「統計学の知識」、「総合判断・産業応用」

さらにそれらは階層的に分類整理された。文献等を参考に各実験分野に必要な知識項目を列挙した上で、有識者委員会、ヒアリングで検討および重要な項目の特定を行なった。

スキル項目が知識項目に対応する形で、初級・中級・上級の3段階に分けて設定され記述された。なお、人材ニーズ調査および有識者委員会では統計学の必要性が指摘され、解析結果の解釈にかかる知識項目として、各実験分野の「解析結果の解釈」に「統計学の知識」が加えられた。

#### (6) 実証講義(研修)による検証

研修では計8回の講義で5つの実験分野すべてをカバーする講義構成とした。いずれの分野でも基礎となる「配列解析」、近年の実験技術の普及による従事者の増加に注目した「発現プロファイル解析」の2分野に時間を重点配分した。受講者に事前・事後自己診断アンケート、各回アンケート、事後の到達度確認テストを行い、調査データを収集、分析した。講義は全回を通して有益との回答を得、好評であった。選択された5つの実験分

野はいずれも有益であるとの回答を受講者より得た。また、講義の順序として「配列解析」が事前に想定した通り基礎であり、最初に学ぶべき分野であることが検証できた。

参加した受講者の職種・専門は多岐にわたり、多方面からのニーズがあることが確認できた。レベル設定をよく検討することにより、より多くの、より多様な人材をこのような育成講座に迎え入れることが可能となり、専任アノテーターとは別に、現在の職や専門をそのまま生かしながらアノテーション能力を習得するバイオ人材を育成できると考えられた。

また、受講者を個別に見ると、5つの実験分野すべてについて理解していないのではなく、特定の実験については比較的知識があるなど、実験分野ごとに理解度のばらつきが見られた。また、それぞれ必要とする実験分野が異なることもあり、実験分野ごとに取捨選択できれば、受講者のニーズに個別に対応することが可能となる。加えて、有識者委員会でも指摘されていたように、5つの分野すべてを習得することは量的に難しいため、スキルスタンダード・カリキュラムが分野ごとに独立の体系を持つことが望ましいと考えられた。

#### (7) アノテーション能力の検討

全8回の講義において、毎回受講者へ行なったアンケート(参考資料参照)では、5つの実験分野いずれの講座も有用であるとの意見が大勢を占めた。受講者の職種・専門分野が多岐にわたることから、5つの実験分野がポストゲノム研究において重要かつ汎用性のあるものであることが示唆された。あるヒアリングでは、日本のバイオ研究・バイオ産業の将来にとっては、このように初級から中級にかけてのスキルを身につけ、もとの職場に戻って業務に役立てるなど、「アノテーション能力」を身につけた人材層が厚くなることは非常に意義深いとの意見が出され、その他のヒアリングおよび有識者委員会でもこれが支持された。そこで上級コースは専任のアノテーターを育成するカリキュラムと位置づけ、初級・中級は上級コースへのステップであると同時に、アノテーション能力を身につけるためのプログラムとして構成を再検証した。

### 3 2 . 人材ニーズ調査

「バイオ産業を支える人材の育成に関する調査(アノテーター)」(以下調査とする)を三井情報開発株式会社と共同で実施し、カリキュラムおよびスキルスタンダードへの反映をおこなった。

#### 3 2 1 調査の概要

本調査は、調査対象企業・研究機関におけるバイオ事業・研究の状況、及びその際に必要となるアノテーター人材の能力・資質等に対する考え等を把握するために、企業及び研究機関を対象として行ったものである。主な調査内容は、アノテーター人材の必要性、不足感、スキル・レベル案の妥当性であるが、さらにスキルスタンダード案から作成した各実験・各項目についての必要性、不足感、育成・獲得方法等についても調査した。100の企業および研究機関へ調査票を発送した結果、37件の回答を得た。

#### 3 2 2 調査結果(全体の結果は参考資料に掲載)

##### (1) スキル項目について

スキル項目の基本構成、および選択した実験分野についての設問では、下に示すようにすべての実験分野・スキル項目について、半数以上の回答者が「極めて必要」もしくは「必要」と回答した。

<設問> 貴研究所内でアノテーション関連業務・研究を担う人材に必要なスキル等に関し、それぞれのスキルの【必要度】として最も近い選択肢をお選びください。  
(個々の設問・回答を、下記テーブルにあわせグループの平均にして表示した)

図表 3-1 アノテーター人材におけるスキルの必要度

	実験の知識・経験		検索・予測ツールの知識・経験		検索・予測結果の解釈	
	極めて必要	あまり必要ではない	極めて必要	あまり必要ではない	極めて必要	あまり必要ではない
	必要	不要	必要	不要	必要	不要
配列解析 (シーケンス)	5.5	5.0	10.7	1.0	10.0	3.0
	15.0	0.5	14.3	0.0	12.7	0.3
発現プロファイル解析 (DNA マイクロアレイ・DNA チップ)	2.0	7.0	5.3	5.7	8.7	4.0
	15.5	1.5	12.7	2.3	11.7	1.7
多型解析(SNP)	3.5	8.0	4.3	4.0	6.0	4.0

	11.5	3.0	16.0	1.7	14.7	1.3
プロテオーム解析 (質量分析)	3.5	8.0	4.0	4.7	5.3	6.3
	12.5	2.0	16.0	1.3	13.7	0.7
タンパク質間ネットワーク (遺伝子ネットワーク)解析	4.0	7.0	3.7	6.0	4.3	6.3
	14.0	1.0	14.7	1.7	14.7	0.7

#### (2) スキルスタンダード基本構造の検討

他の実験分野の追加を指摘する回答は無かったため、大量データを生み出す実験分野としては、配列解析、発現プロファイル解析、多型解析、プロテオーム解析、タンパク質間ネットワーク解析の5分野とした。また、実験の個々のスキル項目についてもすべて過半数の支持を得たことから、原案通り、各実験をアノテーションの流れに沿って実験の理解、データ解析、そして解析結果の解釈の3段階に分割することとした。

#### (3) その他に重要なスキル項目について

調査したスキル項目以外に重要と考えられるスキル項目の指摘が4件あり、「タンパク質・アミノ酸の物理化学的性質」等、アノテーター養成に限らないスキルまたは基礎知識を除き、スキルスタンダード原案中の既にある分類項目内に反映した。

#### (4) スキル・レベルの設定について

育成するアノテーター人材像を想定した初中上級の3段階のスキル・レベル設定については、9割以上の支持があった。わずかながら「妥当ではない」とした回答として、初級ではデータベースの検索・表示に必要なプログラミング言語の必要性の意見があった。中級ではコミュニケーション能力に否定的な意見、また、上級ではアノテーターが研究プロジェクトの中心にはなれないとの意見があった。さらに、上級にはアノテーションされたデータが何に利用できるのかという研究プロジェクトの目的の再考を促す意見もあった。

<設問> 「アノテーター」の理想的な人材像・人材イメージに達するまでのレベルを仮に3段階に分類するとして、私どもでは以下のような分類仮説を立てました。貴社もしくは貴研究所における業務の実態に照らして、この分類は妥当でしょうか。

<分類>

図表 3-2 アノテータースキルレベル分類

レベル・ レベル分類の軸	役割・能力・知識等 (仮説)
-----------------	-------------------

初級者： アノテーション・スタンダード	アノテーション業務のリーダーの指示に従い、あらかじめ決められたアノテーションの方針、データベースのフォーマット、入力仕様などに沿って、アノテーションの作業ができる。
中級者： アノテーション・エキスパート	自立してアノテーションができ、研究者（実験研究者、バイオインフォマティクス研究者）やデータベースシステム開発者と打合せ・討議などのコミュニケーションがとれる。 実験データの信頼性、解析結果の有意性に関して、研究者と討議ができる。
上級者： アノテーション・マネージャー	アノテーションチームのリーダーとして、総合的な視点からプロジェクトの立案、推進、管理を行うことができる。 特定の専門分野に限らず幅広い知識とコミュニケーション能力を持ち、研究者（実験研究者、バイオインフォマティクス研究者）やデータベースシステム開発者などの共同構成員の中で橋渡し役として中心的な役割を果たす。

< 結果 >

図表 3-3 アノテータースキルレベル分類の妥当性について

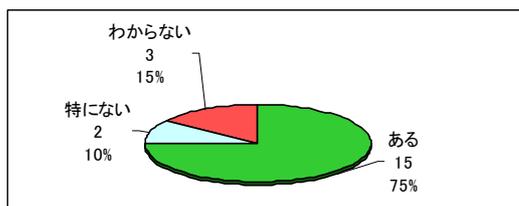
	概ね妥当	妥当ではない
初級者： アノテーション・スタンダード	23 (92.0%)	2 (8.0%)
中級者： アノテーション・エキスパート	22 (91.6%)	2 (8.3%)
上級者： アノテーション・マネージャー	22 (91.6%)	2 (8.3%)

この結果を受け、スキル・レベルにおいては、原案通り初級者は指導者からの指示通りにできる、中級者は自立してできる、上級者はプロジェクトの立案から管理までできるものとした。このレベル設定では「解析結果の解釈」の中の「総合判断・産業応用」は初級には難しいと考えられ、スキルスタンダードおよびカリキュラムでは、どの実験についても「総合判断・産業応用」は中および上級だけができるものとした。さらに、中級と上級間ではレベル差を設けて、上級だけが解析結果からの応用利用を考えることができるものとした。

(5) そのほか調査より得られた知見

現在のデータベースのアノテーションの状況についてたずねたところ、次の通り、回答者のうち実に75%が不備な点や改善点がある、と回答した。

図表 3-4 現在の生物データベースやアノテーションにおいて不備な点や改善が望まれる点



ある	特にない	わからない	無回答
15	2	3	17
(40.5%)	(5.4%)	(8.1%)	(45.9%)

具体的には、多種多様なデータが混在する中で、データベースにあるアノテーション結果が実験で確かめられたものか、配列から予測をした結果なのか、あるいは語彙の統制・制御（オントロジー）が不足しているなど、統一性がないために信頼性が評価できないので、表記法などを含めてフォーマット等の整備を行うことが必要であるというような意見が多くみられた。また、急速に発展した分野であるために、アノテーター自体の経験が不足しているという指摘もあった。

全般的な意見としても、アノテーターの人材育成がきわめて重要である、とする意見が複数みられ、企業や研究機関においてアノテーションを行うことのできる人材が不足しており、この育成が急務であるという状況が伺える。

本調査によって原案が支持され、新興分野の多様な実験を統一して、投影、分類整理可能なスキルスタンダードの基本骨格が決められたことは非常に重要である。

### 3 3 . カリキュラムの作成

#### 3 3 1 概要

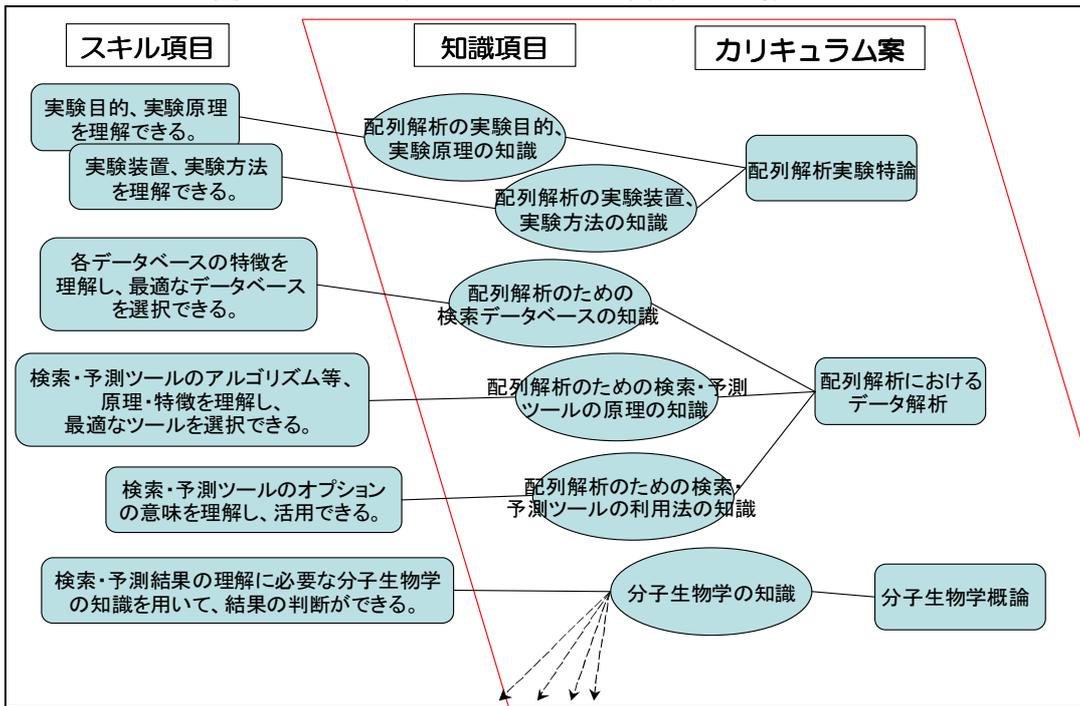
「3 - 1 スキルスタンダードの作成」で述べたように、スキルスタンダードでは業務領域として5つの実験分野に、各実験分野を3段階に、レベルとして初級・中級・上級の3段階とする構造化を行なった。人材ニーズ調査および有識者委員会ではこの基本構造が支持された。特にレベル設定については、人材ニーズ調査において9割以上の支持を得た。これを受けて、カリキュラムの基本構造はスキルスタンダードの構造に倣った。上級レベルの人材育成には初級から段階的にレベルアップを図る必要があり、3段階のレベルのコースを通して、初級・中級レベルの人材を上級レベルにまでステップアップを図ることとし、実験経験者を中心とした生物学・生命科学のバックグラウンドを持つ者を対象に上級者の育成を目指した。一方、様々なバックグラウンドを持つ人材が初級から中級コースに入りやすいように考慮し、科目はすべて5つの実験分野ごとに分けられた。

全てのコースで全レベルの知識項目を含むが、初級では基礎を、中級・上級ではその上のレベルの内容とした。初級から中級へのステップアップは「データ解釈」を自立して行えるよう各実験について科目を増設した。また、中級から上級へのステップアップにおいては「データ解釈」から医療・産業応用に対応できるよう科目を設定した。受講者の持つ実際の経験・知識レベルをスキルスタンダードにて分析し、スタートコースの選択を可能とした。

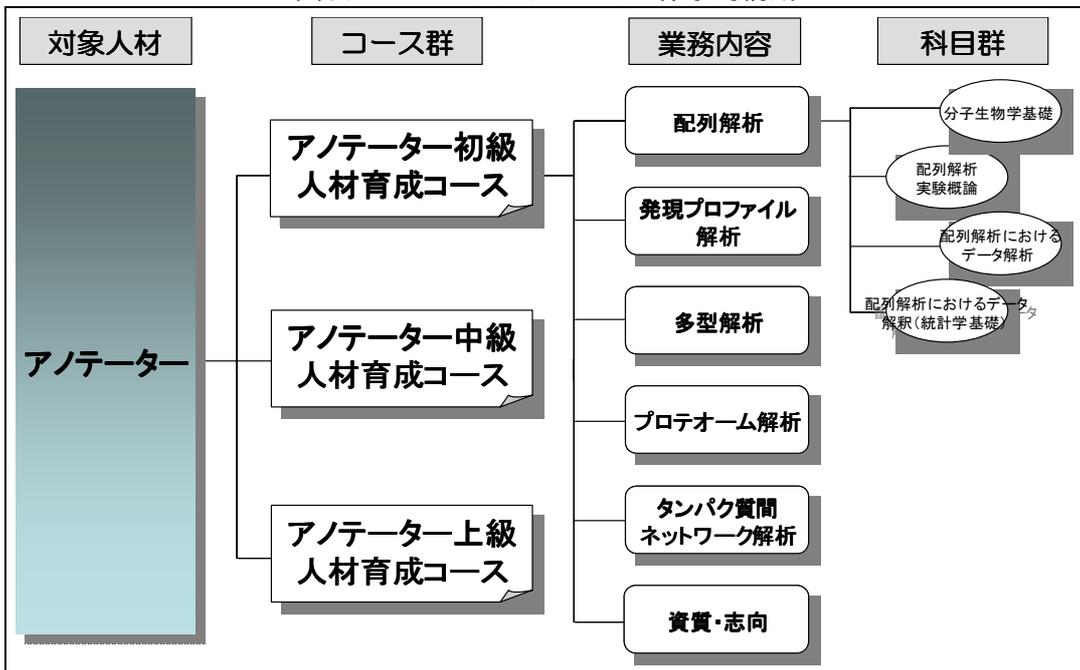
#### 3 3 2 スキルスタンダードからカリキュラムの科目への生成過程

原則、スキル項目は知識項目と一対一の関係で、スキルスタンダードと同様、知識項目は実験の理解、データ解析、解析結果の解釈の3つに整理、集約され、カリキュラムの各科目として設定された。

図表 3-5 知識とカリキュラムの関係(配列解析の例)



図表 3-6 カリキュラムの体系的構成



### 3 3 3 レベル設定

図表 3-7 カリキュラムにおけるレベル設定

コース	育成目標（役割・能力・知識）
アノテーション・スタンダード（アノテーター初級）人材育成コース	<ul style="list-style-type: none"> <li>・アノテーションチームの初級者として、リーダーからの指示を的確に理解してアノテーション作業を行うことができる。</li> <li>・特定の実験分野に限って最低限必要とされる知識を持ち、他の研究者からの指示を理解し、実行できる。</li> <li>・大量データ処理の過程で一連の作業を同じ基準、同じ様式で繰り返し処理できる。</li> </ul>
アノテーション・エキスパート（アノテーター中級）人材育成コース	<ul style="list-style-type: none"> <li>・アノテーションチームの中核として、自立してアノテーションを行うことができる。</li> <li>・特定の専門分野に明るく、初級レベルと上級レベルをつなぐコミュニケーション能力を持ち、チームの中で橋渡しの役割を果たす。</li> <li>・実験のプロセス、複数の解析手法（解析ツール）を把握しており、解析結果を多面的に評価できる。</li> </ul>
アノテーション・マネージャー（アノテーター上級）人材育成コース	<ul style="list-style-type: none"> <li>・アノテーションチームのリーダーとして、総合的な視点からプロジェクトの立案、推進、管理を行うことができる。</li> <li>・特定の専門分野に限らず幅広い知識とコミュニケーション能力を持ち、研究者（実験研究者、バイオインフォマティクス研究者）やデータベースシステム開発者などの共同構成員の中で橋渡し役として中心的な役割を果たす。</li> </ul>

初級コースの目標は指導者からの指示通りにアノテーションできるレベルに留め、「データ解釈」は割愛した。中級コースの目標はアノテーションを自立してできる、上級コースの目標は研究プロジェクトの立案から管理までできるものとした。上級コースの研究プロジェクトの管理等については具体的な事例を伴った項目を考えることは難しかった。

スキルスタンダード案を基に人材ニーズ調査が行われたが、初級・中級・上級の3段階のスキルのレベル設定については、9割以上の支持があった。人材ニーズ調査では、スキルスタンダードの5つの実験分野の各スキル項目について、大差なく「必要」・「不足している」との回答が半数以上あり、スキルスタンダード同様、カリキュラムでは5つの実験分野を網羅する構成・内容にした。

実証講座（研修）では、受講者アンケートにより、いずれの実験分野の講座も有用であるとの意見が大勢を占めた。このことは受講者の職種・専門分野が多岐にわたったことと合わせて、ポストゲノム研究における5つの実験分野の重要性かつ汎用性が支持されたものと考えられた。

### 3 3 4 カリキュラムからのシラバスの作成

シラバスはカリキュラムの構造をそのまま導入し、5つの実験分野ごとに、業務フローに沿った構成とした。シラバスではさらに「解析方法の実習」の項目を加え、各実験分野で必須または代表的な解析ツールを盛り込んだ。レベルについては、各項目を階層で整理してあるので、初級では上の階層を概観して全体像を理解し、中級・上級では、必要な実験分野について、より下層の詳細項目について習得するといった利用方法が考えられる。

### 3 3 5 必要科目の検討

ヒアリング、有識者委員会では、教育科目として統計学の重要性が指摘された。一般的な統計学ではなく、その実験分野に特に必要な統計学を、解析結果の解釈において科目として設定した。初級では「データ解釈（統計学基礎）」とし、中級・上級では標準レベルの統計学として「データ解釈（統計学）」を設けた。ヒアリングでは統計学の個々の項目をすべて教えることは時間的に難しいかもしれないが、例えば、ポワソン分布や尤度といった統計学の項目がこの解析でなぜ必要か、そして、その項目と解析との関係をよく説明することが重要であるとの知見を得た。この観点からは、一般的な統計学ではなく、それぞれの解析に必要な部分を抽出して教えることが重要であり、各実験において異なる内容を想定した科目「データ解釈（統計学）」を設定した。

有識者委員会では、医療分野との関連や創薬目的等アノテーターが行なう業務が何に利用または応用できるのかを考えながら業務自身を遂行する必要が指摘され、実験分野ごとに科目「データ解釈」について、中級では「総合判断」、上級では解析結果からの応用利用をさらに加え「総合判断・産業応用」とした。一方、この科目はレベルが高いため、初級では割愛し重点を実験の理解やデータ解析に置いた科目構成とした。

さらに、有識者委員会において、シラバスでは必須項目をほぼもれなく列挙したことが評価された。一方、これらすべてを一人の受講者が理解するのは難しいとの指摘に対し、本カリキュラム・シラバスは配列解析等5つの実験分野ごとに科目が分けられているため、講義の事業化の際には必要な実験分野科目の取捨選択によって対応が可能であると考えた。また、実学という観点ではシラバスに「解析方法の実習」の項目を加え、各実験分野で必須または代表的な解析ツールを盛り込んだ。個別の実習ツールについても意見が出され、Repeat maskerの実習をシラバスの「配列解析におけるデータ解析」に追加した。その他の追加科目については、情報科学の理論に近いものを除き、指摘された項目をできるだけカリキュラム・シラバスに盛り込むこととした。

また、実証講座（研修）は「データ解析」に必要な解析ツールの実習を中心に行われ、各回の講義で紹介された解析ツールはシラバスを中心に反映されたが、その汎用性が高いと判断されたものはスキルスタンダードに追加した。

### 3 3 6 カリキュラムの実用性及び汎用性

中間報告会では、実用性や汎用性という視点でスキルスタンダード、カリキュラムを修正することを求められた。多くは「スキルスタンダードの作成」および「次年度以降の展開方針」の項で検討した。シラバスでは、それぞれの実験分野に必須の解析ツールを文献調査等で列挙し、有識者委員会・ヒアリングでの検討、および実証講座での知見から精選したものを実習する項を盛り込み、より最新かつ実用的なものとした。

### 3 3 7 次年度以降のカリキュラム・シラバスの活用

有識者委員会では、このようなカリキュラムによる教育を継続して行なうべきであり、また、そのような人材育成の専門家を含むチームづくり、センターを設置するべきとの意見があった。そのようなプログラムに的確に合致するカリキュラムとなるように、5つの実験分野という明確な入り口、そして具体的なメニューとしての3つのコース分けによって本カリキュラムを最大限に活用することが期待できる。

### 3 4 . 実証プログラムの実施

大量データを扱う代表的な実験について真に有効なアノテーションを行うために必要とされる 3 つの能力、すなわち 生データの信頼性も評価できる実験自体の理解、 検索・予測プログラムの知識・活用といったデータの解析力、そして 分子生物学と情報科学に支えられた解析結果の解釈力を身につけるため、単なる知識の習得を目的にした講義だけでなく、演習を積極的に取り入れた。最新の解析ツールを用いた実習を核に、データ解析だけでなく実験の理解、解析結果の解釈まで試みた。講義がカバーする実験分野は、1) 配列解析、2) 発現プロファイル解析、3) 多型解析、4) プロテオーム解析、5) タンパク質間ネットワーク解析の 5 分野である。

#### 3 4 1 実証プログラムの概要

図表 3-8 実証プログラムの概要

実証研修テーマ：	アノテーター中級・上級人材育成コースの実施
日程・時間：	平成 16 年 10 月 13 日(水)～12 月 15 日(水)毎週水曜日 計 8 日間
	13:30～17:00
方法：	座学による講義・実習、アンケートおよび到達度テスト
対象者：	バイオを専門とする学士、修士、ポスドク、企業研究者等、バイオに関して一通りの知識を持つ者(ただし実務未経験者も可)として募集した。
研修成果の測定方法・スキルスタンダード、カリキュラムへの反映方法	
測定方法：	受講前と受講後のアンケートおよび受講後の到達度テストを実施し、習熟度を測定する。
反映方法：	受講効果測定及びその結果に対する委員会での検討を通じて、当該業務領域の上級までのスキルスタンダードを検討・確立する。

平成 16 年 10 月 13 日(水)～12 月 15 日(水)の原則毎週水曜日 13:30～17:00(講義時間 3 時間)に計 8 回、合計 24 時間の実証講座を設定、実施した。講義はコンピューターを利用した実習を中心とし、座学の形式で行われた。限られた講義時間の中で 5 つの実験分野のすべてをカバーしたが、実験機器の普及等から従事人口が多いことが予想される配列解析、発現プロファイル解析に時間を重点的に配分した。モニター受講者は 17 名で、受講者の主な所属は、製薬・化学メーカー、治験担当者(薬剤師)、医師、獣医、専門学校教員、人材派遣業、公的機関研究者、IT 通信、ソフトウェア開発であった。

図表 3-9 実証講義の日程

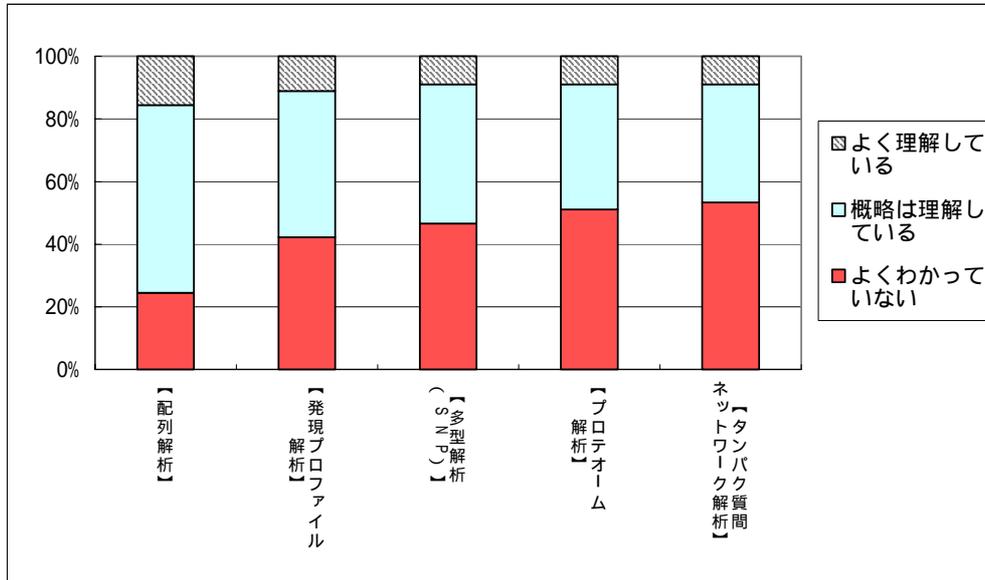
講義日程		講義内容	講師名	所属
10 月 13 日	水	配列解析 1	大城戸 利久	日本 DNA データバンク
10 月 20 日	水	配列解析 2	小菅 武英	日本 DNA データバンク

講義日程		講義内容	講師名	所属
10月27日	水	配列解析3	田中 剛	生命情報・DDBJ 研究センター
11月10日	水	発現プロファイル解析1	荻島 創一	東京医科歯科大学
11月17日	水	発現プロファイル解析2	荻島 創一	東京医科歯科大学
11月24日	水	発現プロファイル解析3	荻島 創一	東京医科歯科大学
12月4日	土	多型解析(SNP)	今西 規	生物情報解析研究センター
12月15日	水	プロテオーム解析とタンパク質相互間解析	櫻井 仁美	生命情報・DDBJ 研究センター

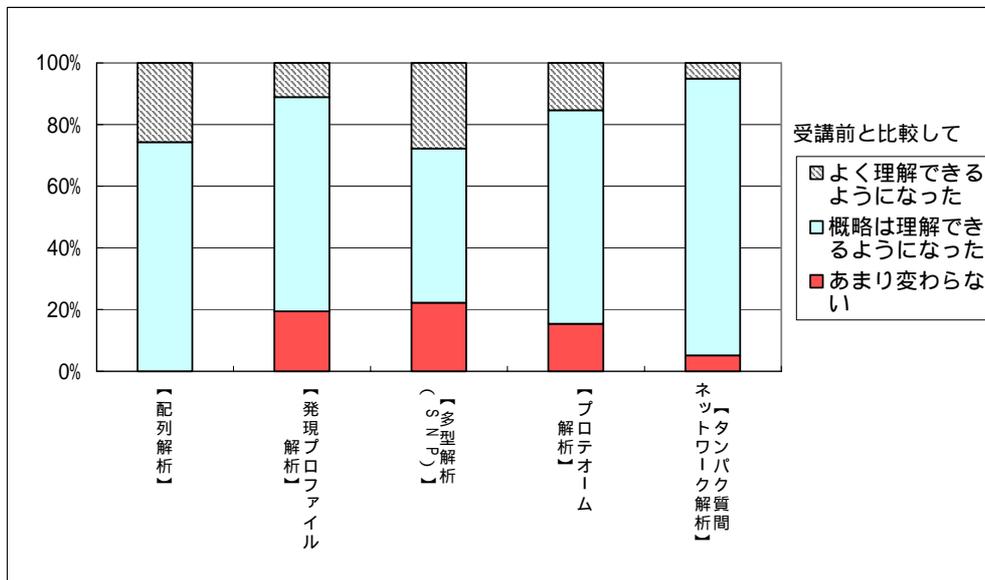
(1) 自己評価アンケート

モニター受講者の意識・レベルを調べるため、実証講座開始前、および全講座終了後に自己評価のアンケートを実施した。

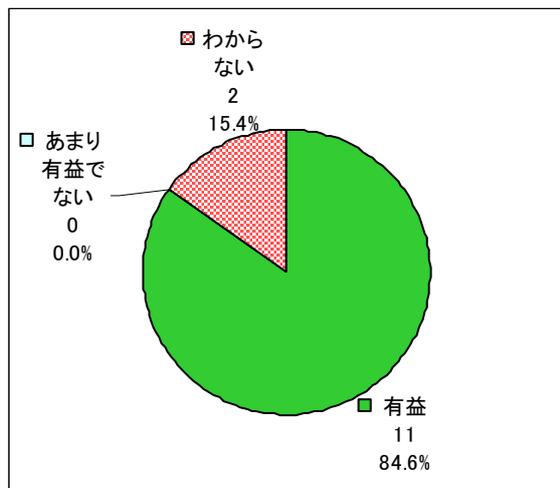
図表 3-10 事前自己評価アンケートの集計結果



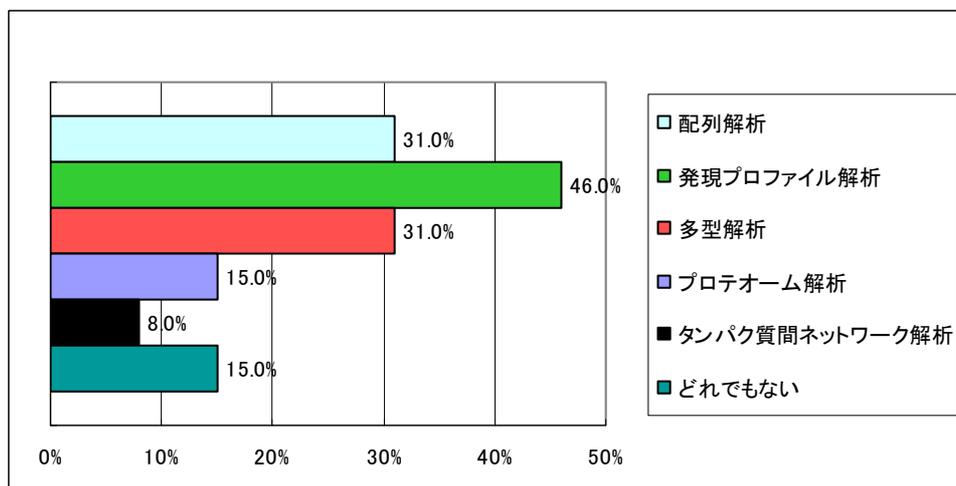
図表 3-11 事後自己評価アンケートの集計結果



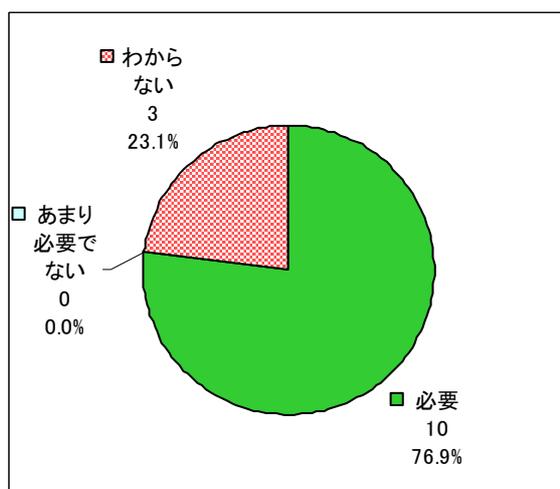
図表 3-1 2 本講座は受講者自身の今後の研究・ビジネス活動にとって有益でしたか



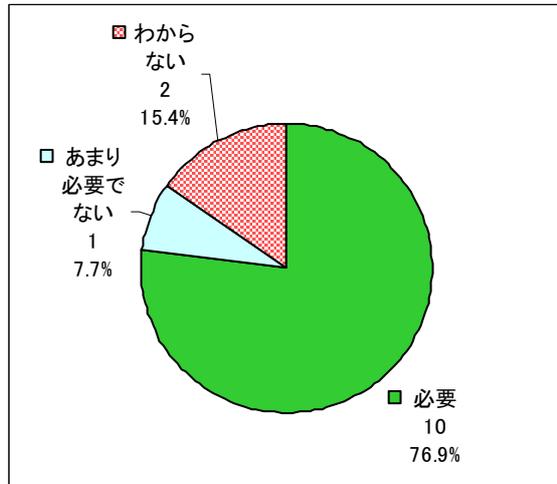
図表 3-1 3 最も有益であった講義分野（複数回答）(N=13)



図表 3-1 4 アノテーター人材の必要性



図表 3-15 今後の研修の必要性



図表 3-16 受講者の経歴について(人)

バイオ経験年数		IT 経験年数	
~1年	1	~1年	4
~5年	2	~5年	3
~10年	3	~10年	1
~15年	5	~15年	1
~20年	1	~20年	0
20年~	1	20年~	0

受講者の主な経歴：

バイオインフォマティクス(4)、教員(2)、SNP、疾患関連遺伝子探索、遺伝子発現制御メカニズム解析、組み換えニワトリ用のワクチンの開発、ウイルス改変、外来遺伝子導入、創薬、二次代謝産物の生合成酵素の構造、機能解析、バイオプロセス関連のサポート業務

(2) 実証講義(研修)結果およびスキルスタンダード・カリキュラムへの反映

講義前の事前自己評価アンケートによれば、5つの各実験分野をよく理解している受講者は9~16%、よくわかっていない受講者は24~53%であった。これに対し、実証講座の終了後に毎行行ったアンケートでは、講義内容が参考になったと答えた受講者がほとんどであった。アノテーターに必要とされる5つの実験分野について、受講前の受講者の理解は十分でなかった。事前自己評価アンケートからは直接見えないが、5つの実験分野について個々の受講者の理解度にはばらつきがあった。また、講座終了後のアンケートの「最も有益であった講義分野」でも各受講者が必要とする実験分野は多様であった。これらのことは、スキルスタンダード・カリキュラムを5つの実験分野に分解することにより、それらを受講者が取捨選択できる効率的な育成が可能になると考えられる。

受講者として多彩な所属からの参加があり、その年齢もおよそ30代から50代と広い範囲にわたると推測された。また、人材ニーズ調査が調査対象とした指導者層よりも、ポストゲノム研究に直接関わる実務担当者が多いと推定されることは興味深い。そのほとんどが講義内容について有益であったと答えていることから、スキルスタンダードの基本構造としての5つの実験分野が多様な分野の人材に対応できるスタンダードとして適切であることが裏付けられた。また、今回は業務フロー3段階中の「データ解析」・「解析結果の解釈」を中心に講義が行われたが、各回とも行われた実習では、研究室で実際にアノテーションに使われたデータが用いられ、そのデータの由来やエラーが含まれる可能性等、

「実験の理解」を意識した講義が行われた。「実験の理解」をさらに深めるために実験機器を用いた実験室での研修を別途行うことも、可能であり、業務フローを3段階に分けたことによって可能である。

各回の講義内容から、スキル項目またはカリキュラムの一項目として必要と考えられたものについては、それぞれの分類項目内に追加した。特に最新の解析ツールは実用に耐えるものとして列挙した。

受講後の到達度テストでは各実験分野とも7~9割の理解が確認できたが、筆記によるテストであり、解析ツールを使いこなせるかというパソコンを用いた実技については別途測定する必要がある。また、アノテーターは未知の遺伝子・タンパク質の機能を推定するという「解析結果の解釈」を行う必要があり、このスキルの理解度を測定することは非常に難しい。

講座終了後のアンケートで、毎回の講座がやさしすぎると答えた受講者はほとんどいなかった。また、アノテーター講座を必要として参加した受講者の専門・所属は多岐にわたっていた。このようなことから、様々な分野からの参加をやすくし、多様な人材がこの業務に参加するためにも間口を広くすることが必要と考えられ、コースとカリキュラムの設計においては、初級コースについては基礎的なものに絞ることとした。多様かつより多くの人材を初級コースに導き入れ、初級からのレベルアップによって、最終目標である中級・上級レベルの人材を育成することが可能となる。

さらにアンケートでは講義時間の延長や次回の講座開催への期待が多数あり、講座を初めとする教育が依然として必要とされており、アノテーター育成のニーズは確実に存在していると考えられる。



## 第4章 スキルスタンダード

---

### 4 1 . スキル項目とスキルレベル

#### 4 1 1 育成対象とする人材像

本スキルスタンダードは、以下の人材を対象として想定し作成されている。

- 1) 専任アノテーターという高度バイオ人材の育成  
最終目標であるトップレベルの人材であり、上級レベルとして想定。
- 2) 兼任アノテーターという企業内・研究所内においてアノテーションをおこなう人材の育成  
社内データベースを活用する企業でのアノテーション業務を遂行する人材を中級～上級レベルとして想定。
- 3) 生物学・生命科学系人材が必要とするバイオインフォマティクス基本技能の習得  
アノテーション技能の習得が、実験から産出される生物データを解析、解釈するバイオインフォマティクス技術を含んでおり、その基本技能の習得を初級～中級レベルとして想定。

#### 4 1 2 スキルスタンダードの基本構成

スキルスタンダードを作成するにあたり、アノテーターは大量の生物データが生み出されるところに必要であるという人材像を設定しており、スキルスタンダードの最も基本的な分類を実験分野別としている。実験分野の選定にあたってはデータ処理にコンピューターが不可欠となるような大量のデータを生み出し、ポストゲノム研究の中で重要かつ代表的な実験分野として以下を選定した。

##### 【実験分野】

- 1) 配列解析 (シーケンス)
- 2) 発現プロファイル解析 (DNA マイクロアレイ・DNA チップ)
- 3) 多型解析 (SNP)
- 4) プロテオーム解析 (質量分析)
- 5) タンパク質間ネットワーク (遺伝子ネットワーク) 解析
- 6) 資質・志向

また、各実験分野は共通して、データの流れに沿って以下の3段階の業務フローに分類されている。

##### 【業務フロー】

- 1) 実験内容の理解

- 2) データの解析
- 3) 解析結果の解釈

さらにスキル項目は以下のレベルで3段階に分けられている。

#### 【スキルレベル】

- 1) 初級レベル(アノテーション・スタンダード)
  - ・ アノテーション業務のリーダーの指示に従い、あらかじめ決められたアノテーションの方針、データベースのフォーマット、入力仕様などに沿って、アノテーションの作業ができる。
- 2) 中級レベル(アノテーション・エキスパート)
  - ・ 自立してアノテーションができ、複数の解析結果を選択、統合して解釈できる。
  - ・ 研究者(実験研究者、バイオインフォマティクス研究者)やデータベースシステム開発者と打合せ・討議などのコミュニケーションがとれる。
  - ・ 実験データの信頼性、解析結果の有意性に関して、研究者と討議ができる。
- 3) 上級レベル(アノテーション・マネージャー)
  - ・ アノテーションチームのリーダーとして、総合的な視点からプロジェクトの立案、推進、管理を行うことができる。
  - ・ 特定の専門分野に限らず幅広い知識とコミュニケーション能力を持ち、研究者(実験研究者、バイオインフォマティクス研究者)やデータベースシステム開発者などの共同構成員の中で橋渡し役として中心的な役割を果たす。

以上の基本構造に従って、各分類に対応する知識項目およびスキル項目で不可欠なものを精選し、階層的に整理展開してある。

これにより具体的な実験分野を選択し、各人のレベルに合わせ、そして業務フローの必要な箇所から育成を開始することが可能となり、初級から段階的に上級の専任アノテーターを育成するだけでなく、現在の職種・専門を維持したままアノテーション能力を身に付けるコースも編成できる構成となっている。

## 4 2 . スキルスタンダード

### 4 2 1 配列解析 (シーケンス)

業務内容	知識項目	スキル項目		
		初級	中級	上級
		(但し、初級・中級・上級の違いは理解すべき項目や理解すべきレベルによって差別化する。)		
実験の理解	配列解析の実験目的、実験原理の知識	実験目的、実験原理を理解できる。	実験目的、実験原理を十分理解できる。	実験目的、実験原理を十分理解できる。
	<ul style="list-style-type: none"> <li>・ クローニング               <ul style="list-style-type: none"> <li>➢ BAC クローン、cDNA クローン</li> <li>➢ cDNA ライブラリー</li> <li>➢ ベクター</li> <li>➢ 全長 cDNA クローニング法</li> </ul> </li> <li>・ ショットガンシーケンス</li> <li>・ Assemble</li> <li>・ シーケンサー               <ul style="list-style-type: none"> <li>➢ 電気泳動</li> </ul> </li> </ul>	実験目的、実験原理の概略を理解できる。	実験目的、実験原理を十分理解できる。	実験目的、実験原理を十分理解できる。
	配列解析の実験装置、実験方法の知識	実験装置、実験方法を理解できる。	実験装置、実験方法を理解できる。	実験装置を用いて対象データを得るプロセスを十分理解しており、データの信頼性について評価できる。
	<ul style="list-style-type: none"> <li>・ シーケンサー</li> <li>・ Base caller プログラム               <ul style="list-style-type: none"> <li>➢ Phred</li> </ul> </li> <li>・ Assemble プログラム               <ul style="list-style-type: none"> <li>➢ Phrap</li> </ul> </li> </ul>	実験装置、実験方法の概略を理解できる。	実験装置、実験方法を理解できる。	実験装置を用いて対象データを得るプロセスを十分理解しており、データの信頼性について評価できる。

検索・予測ツールを用いたデータ解析	配列解析のための検索データベースの知識	各データベースの特徴を理解し、最適なデータベースを選択できる。		
	<ul style="list-style-type: none"> <li>複数のデータベースの知識・比較</li> <li>データ量</li> <li>データの冗長性</li> <li>網羅度</li> <li>更新間隔</li> </ul>	代表的なデータベースの特徴を理解できる。	十分な種類のデータベースの特徴を理解しており、それらから最適なデータベースを選択できる。	十分な種類のデータベースの特徴を理解しており、それらから最適なデータベースを選択できる。
	配列解析のための検索・予測ツールの原理の知識	検索・予測ツールのアルゴリズム等、原理・特徴を理解し、最適なツールを選択できる。		
	<ul style="list-style-type: none"> <li>ホモロジー検索</li> <li>遺伝子予測 <ul style="list-style-type: none"> <li>遺伝子構造</li> <li>隠れマルコフモデル (HMM)</li> <li>比較ゲノム解析</li> </ul> </li> <li>モチーフ検索</li> <li>プロファイル (ドメイン) 検索 <ul style="list-style-type: none"> <li>細胞内局在予測 <ul style="list-style-type: none"> <li>シグナルペプチド</li> </ul> </li> <li>膜貫通領域</li> </ul> </li> <li>二次構造予測</li> <li>立体構造予測</li> </ul>	検索・予測ツールのアルゴリズム等、原理・特徴の概略を理解できる。	検索・予測ツールのアルゴリズム等、原理・特徴の概略を理解しており、それらから代表的なツールを選択できる。	各ツールの原理・特徴を理解しており、それらから最適なツールを選択できる。
	配列解析のための検索・予測ツールの利用法の知識	検索・予測ツールのオプションの意味を理解し、活用できる。		
	<ul style="list-style-type: none"> <li>オプションの最適化</li> </ul>	指定されたオプションを利用できる。	ツールのオプションの概要を理解し、基本的なオプションを利用できる。	ツールの各オプションを十分に理解し、基本的なオプションを組み合わせ活用できる。
検索・予測結果の解釈	分子生物学の知識	検索・予測結果の理解に必要な分子生物学の知識を用いて、結果の判断ができる。		
	<ul style="list-style-type: none"> <li>分子生物学用語の知識 <ul style="list-style-type: none"> <li>マッピング</li> </ul> </li> <li>文献読解</li> <li>Gene ontology</li> <li>.</li> </ul>	検索・予測結果中の基本的な専門用語が理解できる。	検索・予測結果中の専門用語が理解でき、内容を取捨選択し、総合的な判断ができる。関連する文献内容が理解できる。	検索・予測結果中の専門用語が理解でき、内容を取捨選択し、総合的な判断ができる。関連する文献内容が理解できる。

	配列解析のための統計学の知識	検索・予測結果の数値から結果の信頼性（統計的有意性）が判断できる。		
	<ul style="list-style-type: none"> <li>・ 確率分布 <ul style="list-style-type: none"> <li>➢ E-value</li> </ul> </li> <li>・ 標準偏差 <ul style="list-style-type: none"> <li>➢ Z score</li> </ul> </li> <li>・ 尤度</li> </ul>	検索・予測結果の判断に必要な統計用語の概略を理解できる。	検索・予測結果の数値から結果の信頼性（統計的有意性）が判断できる。	統計学の基礎を理解し、検索・予測結果の数値から結果の信頼性（統計的有意性）が判断できる。
	配列解析のための総合判断・産業応用の知識	検索・予測結果を総合的に判断し、産業への応用を検討できる。		
	<ul style="list-style-type: none"> <li>・ 機能推定</li> <li>・ 検索・予測結果の比較・検討</li> <li>・ 医療及び産業応用に関する知識</li> </ul>		検索・予測結果から配列の機能を推定できる。 オプション変更や他のツールでの再検索により、結果を比較検討できる。	検索・予測結果から配列の機能を推定できる。 オプション変更や他のツールでの再検索により、結果を比較検討できる。 疾患関連遺伝子等、機能遺伝子の産業への応用を検討できる。

4 2 2 発現プロファイル解析 (DNA マイクロアレイ・DNA チップ)

業務内容	知識項目	スキル項目		
		初級	中級	上級
		(但し、初級・中級・上級の違いは理解すべき項目や理解すべきレベルによって差別化する。)		
実験の理解	発現プロファイル解析の実験目的、実験原理の知識	実験目的、実験原理を理解できる。		
	<ul style="list-style-type: none"> <li>・ 遺伝子発現</li> <li>・ ハイブリダイゼーション</li> <li>・ 蛍光ラベリング</li> </ul>	実験目的、実験原理の概略を理解できる。	実験目的、実験原理を十分理解できる。	実験目的、実験原理を十分理解できる。
実験の理解	発現プロファイル解析の実験装置、実験方法の知識	実験装置、実験方法を理解できる。		
	<ul style="list-style-type: none"> <li>・ マイクロアレイ、DNA チップ</li> <li>・ 調整試料 <ul style="list-style-type: none"> <li>➢ cDNA ライブラリー</li> <li>➢ RNA 抽出</li> <li>➢ PCR</li> </ul> </li> <li>・ 蛍光色素 (Cy3, Cy5)</li> <li>・ スキャナー</li> <li>・ アレイ / チップ上の遺伝子の既知情報</li> </ul>	実験装置、実験方法の概略を理解できる。 DNA チップのロット、使用した読み取りスキャナー番号、実験者等、データベース入力に必要な項目を把握している。	実験装置、実験方法を理解できる。 データベース入力に必要な項目に加え、DNA チップ上の遺伝子の種類、使用された試料、実験条件等を把握している。	実験装置を用いて対象データを得るプロセスを十分理解しており、データの信頼性について評価できる。
解析ツールを用いる	発現プロファイル解析のための解析ツールの知識	スポットの蛍光強度の正規化方法及びクラスター解析の特徴を理解し、最適なツールを選択できる。		

たデータ解析	<ul style="list-style-type: none"> <li>・ 各ツールの知識・比較 <ul style="list-style-type: none"> <li>➢ ScanAlyze</li> <li>➢ PRIM</li> <li>➢ Cluster</li> <li>➢ TreeView</li> <li>➢ GeneSpring</li> <li>➢ Spotfire</li> </ul> </li> </ul>	正規化方法及びクラスター解析の概略を理解できる。	複数の正規化方法及びクラスター解析の特徴を理解できる。	複数の正規化方法及びクラスター解析の特徴を理解できる。
発現プロファイル解析のための解析ツールの原理の知識		スポットの蛍光強度の正規化方法及びクラスター解析の原理を理解し、最適なツールを選択できる。		
	<ul style="list-style-type: none"> <li>・ 正規化</li> <li>・ クラスター解析</li> </ul>	正規化方法及びクラスター解析の原理の概略を理解できる。	複数の正規化方法及びクラスター解析の原理を理解できる。	複数の正規化方法及びクラスター解析の原理を理解できる。
発現プロファイル解析のための解析ツールの利用法の知識		解析ツールを活用できる。		
	<ul style="list-style-type: none"> <li>・ 蛍光強度の定量化 <ul style="list-style-type: none"> <li>➢ 正規化</li> <li>➢ バックグラウンド評価</li> <li>➢ はずれ値検出</li> </ul> </li> <li>・ クラスター解析 <ul style="list-style-type: none"> <li>➢ 相関</li> </ul> </li> </ul>	指示に従って、バックグラウンドを見積もることによりはずれ値を検出できる。指定された正規化方法を用いてスポットの蛍光強度を定量化できる。指定された手法のクラスター解析ができる。	バックグラウンドを見積もることによりはずれ値を検出できる。正規化及びクラスター解析の各手法の比較ができ、最適の方法を選択できる。	正規化及びクラスター解析の各手法の比較ができ、最適の方法を選択できる。他のデータとの比較によりスポットの蛍光強度の再現性が検討できる。
解析結果の解釈	分子生物学の知識	相関する遺伝子群に共通の機能を見つけ出すのに必要な分子生物学の知識を用いて、結果の判断ができる。		

	<ul style="list-style-type: none"> <li>・ 分子生物学用語の知識</li> <li>・ 文献読解</li> <li>・ Gene ontology</li> <li>・ アレイ / チップ上の遺伝子の既知情報 <ul style="list-style-type: none"> <li>➢ 機能情報</li> <li>➢ 配列類似性</li> <li>➢ 関連文献</li> </ul> </li> </ul>	<p>解析結果中の基本的な専門用語が理解できる。</p>	<p>解析結果中の専門用語が理解でき、内容を取捨選択し、総合的な判断ができる。 関連する文献内容が理解できる。</p>	<p>解析結果中の専門用語が理解でき、内容を取捨選択し、総合的な判断ができる。 関連する文献内容が理解できる。</p>
	<p>発現プロファイル解析のための統計学の知識</p>	<p>適切なクラスター解析と閾値の選択により遺伝子をグループ（クラスター）化して、発現パターンが相関・逆相関しているものを見つけるための統計的有意性を評価できる。</p>		
	<ul style="list-style-type: none"> <li>・ クラスター解析 <ul style="list-style-type: none"> <li>➢ 相関</li> </ul> </li> <li>・ 多変量解析</li> </ul>	<p>正規化及びクラスター解析に必要な統計用語の概略を理解できる。</p>	<p>統計学の基礎を理解できる。 正規化及びクラスター解析に用いた閾値の統計的有意性が判断できる。</p>	<p>統計学の基礎を理解できる。 解析手法及び閾値から結果の信頼性（統計的有意性）が判断できる。 閾値の変更や他の手法で解析を行い、結果を比較検討できる。</p>
	<p>発現プロファイル解析のための総合判断・産業応用の知識</p>	<p>解析結果を総合的に判断し、産業への応用を検討できる。</p>		

	<ul style="list-style-type: none"> <li>・ 機能推定 <ul style="list-style-type: none"> <li>➢ 組織特異的発現</li> </ul> </li> <li>・ 遺伝子ネットワーク</li> <li>・ システムバイオロジー</li> <li>・ 解析結果の比較・検討</li> <li>・ 医療及び産業応用に関する知識</li> </ul>		<p>解析結果から遺伝子の機能を推定できる。 オプション変更や他方法でのクラスタリングにより、結果を比較検討できる。</p>	<p>解析結果から遺伝子の機能を推定できる。 オプション変更や他方法でのクラスタリングにより、結果を比較検討できる。 疾患関連遺伝子等、機能遺伝子の産業への応用を検討できる。</p>
--	--	--	--	---

4 2 3 多型解析 (SNP)

業務内容	知識項目	スキル項目		
		初級	中級	上級
		(但し、初級・中級・上級の違いは理解すべき項目や理解すべきレベルによって差別化する。)		
実験の理解	多型解析の実験目的、実験原理の知識	実験目的、実験原理を理解できる。		
	<ul style="list-style-type: none"> <li>・ シーケンス</li> <li>・ Assemble</li> <li>・ Invader 法</li> <li>・ ICAM 法</li> </ul>	実験目的、実験原理の概略を理解できる。	実験目的、実験原理を十分理解できる。	実験目的、実験原理を十分理解できる。
	多型解析の実験装置、実験方法の知識	実験装置、実験方法を理解できる。		
	<ul style="list-style-type: none"> <li>・ シーケンサー</li> <li>・ Assemble プログラム</li> <li>・ Invader 法</li> <li>・ ICAM 法</li> </ul>	実験装置、実験方法の概略を理解できる。	実験装置、実験方法を理解できる。	実験装置を用いて対象データを得るプロセスを十分理解しており、配列決定された泳動像を参照して配列決定の精度を確認できる。
検索・予測ツールを用いたデータ解析	多型解析のための検索データベースの知識	SNP を含む配列の検索に必要な各データベースの特徴を理解し、最適なデータベースを選択できる。		
	<ul style="list-style-type: none"> <li>・ 複数のデータベースの知識・比較                             <ul style="list-style-type: none"> <li>➢ dbSNP</li> <li>➢ JSNP</li> </ul> </li> <li>・ データ量</li> <li>・ データの冗長性</li> <li>・ 網羅度</li> <li>・ 更新間隔</li> </ul>	代表的なデータベースの特徴を理解できる。	十分な種類のデータベースの特徴を理解しており、それらから最適なデータベースを選択できる。	十分な種類のデータベースの特徴を理解しており、それらから最適なデータベースを選択できる。
	多型解析のための検索・予測ツールの原理の知識	SNP を含む配列の機能の解析に必要な検索・予測ツールの原理を理解し、最適なツールを選択できる。		

	<ul style="list-style-type: none"> <li>ホモロジー検索</li> <li>マルチプルアライメント</li> <li>遺伝子予測</li> <li>Base caller</li> <li>Assemble</li> </ul>	検索・予測ツールのアルゴリズム等、原理・特徴の概略を理解できる。	検索・予測ツールのアルゴリズム等、原理・特徴の概略を理解できる。	検索・予測ツールのアルゴリズム等、原理・特徴を理解できる。
	多型解析のための検索・予測ツールの利用法の知識	SNPを含む配列の機能の解析に必要な検索・予測ツールを活用できる。		
	<ul style="list-style-type: none"> <li>Base caller プログラム（配列精度の評価） <ul style="list-style-type: none"> <li>Phred</li> </ul> </li> <li>Assemble プログラム <ul style="list-style-type: none"> <li>Phrap</li> </ul> </li> <li>ホモロジー検索</li> <li>遺伝子予測 <ul style="list-style-type: none"> <li>遺伝子構造</li> </ul> </li> </ul>	ホモロジー検索により SNP を含む領域に一致する配列を検索できる。Assemble プログラムから SNP を同定できる。	ホモロジー検索により SNP を含む領域に一致する配列を検索できる。Assemble プログラムから SNP を同定できる。Base caller プログラムを使い、配列精度を評価できる。	ホモロジー検索により SNP を含む領域に一致する配列を検索できる。Assemble プログラムから SNP を同定できる。Base caller プログラムを使い、配列精度を評価できる。
検索・予測結果の解釈	分子生物学の知識	検索・予測結果を分子生物学の知識で解釈できる。		
	<ul style="list-style-type: none"> <li>分子生物学用語の知識</li> <li>文献読解</li> <li>Gene ontology</li> <li>変異 <ul style="list-style-type: none"> <li>アミノ酸残基置換</li> <li>コドン</li> </ul> </li> </ul>	検索・予測結果中の基本的な専門用語が理解できる。検索・予測結果から SNP を含む配列の機能を推定できる。	検索・予測結果中の専門用語が理解でき、内容を取捨選択し、総合的な判断ができる。検索・予測結果から SNP を含む配列の機能を推定できる。関連する文献内容が理解できる。	検索・予測結果中の専門用語が理解でき、内容を取捨選択し、総合的な判断ができる。検索・予測結果から SNP を含む配列の機能を推定できる。関連する文献内容が理解できる。
	多型解析のための統計学の知識	検索・予測結果の数値から結果の信頼性（統計的有意性）が判断できる。		
	<ul style="list-style-type: none"> <li>遺伝統計学 <ul style="list-style-type: none"> <li>連鎖解析 <ul style="list-style-type: none"> <li>パラメトリック</li> <li>ノンパラメトリック</li> <li>連鎖不平衡</li> </ul> </li> <li>QTL 解析</li> <li>ハプロタイプ解析</li> </ul> </li> </ul>	検索・予測結果の判断に必要な統計用語の概略を理解できる。	検索・予測結果の数値から結果の信頼性（統計的有意性）が判断できる。	統計学の基礎を理解し、検索・予測結果の数値から結果の信頼性（統計的有意性）が判断できる。
	多型解析のための総合判断・産業応用の知識	解析結果を総合的に判断し、産業への応用を検討できる。		

	<ul style="list-style-type: none"> <li>・ 医療及び産業応用に関する知識 <ul style="list-style-type: none"> <li>➢ 単一遺伝子疾患</li> <li>➢ 多因子疾患</li> </ul> </li> <li>・ 薬剤応答 <ul style="list-style-type: none"> <li>➢ テーラーメイド医療</li> </ul> </li> </ul>		<p>解析結果から SNP と疾患または薬剤応答性との関連を推定できる。</p> <p>オプション変更や他集団との比較により、結果を比較検討できる。</p>	<p>解析結果から SNP と疾患または薬剤応答性との関連を推定できる。</p> <p>オプション変更や他集団との比較により、結果を比較検討できる。</p>
--	---	--	--	--

4 2 4 プロテオーム解析（質量分析）

業務内容	知識項目	スキル項目		
		初級	中級	上級
		(但し、初級・中級・上級の違いは理解すべき項目や理解すべきレベルによって差別化する。)		
実験の理解	プロテオーム解析の実験目的、実験原理の知識	実験目的、実験原理を理解できる。		
	<ul style="list-style-type: none"> <li>・ 二次元ポリアクリルアミド電気泳動 (2D-PAGE)</li> <li>・ プロテアーゼ消化</li> <li>・ 液体クロマトグラフィー (LC)</li> <li>・ イオン化法</li> <li>・ 質量分析法</li> </ul>	実験目的、実験原理の概略を理解できる。	実験目的、実験原理を十分理解できる。	実験目的、実験原理を十分理解できる。
	プロテオーム解析の実験装置、実験方法の知識	実験装置、実験方法を理解できる。		
	<ul style="list-style-type: none"> <li>・ イオン化法 <ul style="list-style-type: none"> <li>➢ エレクトロスプレーイオン化 (ESI) 法</li> <li>➢ マトリックス支援レーザー脱離イオン化 (MALDI) 法</li> </ul> </li> <li>・ 質量分析法 <ul style="list-style-type: none"> <li>➢ 飛行時間型質量分析計 (TOF - MS)</li> <li>➢ 四重極型質量分析計 (Q - MS)</li> <li>➢ イオントラップ型質量分析計 (IT - MS)</li> </ul> </li> <li>・ プロテアーゼの基質特異性</li> </ul>	実験装置、実験方法の概略を理解できる。	実験装置、実験方法を理解できる。	実験装置を用いて対象データを得るプロセスを十分理解しており、データの信頼性について評価できる。 MALDI-TOF、LC-MS 等、各実験方法の特徴を把握している。
解析ツールを用いたデータ解析	プロテオーム解析のための検索データベースの知識	最新かつ最大のアミノ酸配列データベースを把握しており、検索に利用できる。		
	<ul style="list-style-type: none"> <li>・ 複数のデータベースの知識・比較</li> <li>・ データ量</li> <li>・ データの冗長性</li> <li>・ 網羅度</li> <li>・ 更新間隔</li> </ul>	指定されたデータベースの概要を理解できる。	代表的なデータベースの特徴を理解できる。	十分な種類のデータベースの特徴を理解しており、それらから最適なデータベースを選択できる。
	プロテオーム解析のための解析ツールの原理の知識	解析ツールのアルゴリズム等、原理・特徴を理解し、最適なツールを選択できる。		

	<ul style="list-style-type: none"> <li>PMF (peptide mass finger printing) 法</li> <li>MS/MS ( tandem MS ) 法</li> </ul>	解析ツールのアルゴリズム等、原理・特徴の概略を理解できる。	解析ツールのアルゴリズム等、原理・特徴の概略を理解できる。	解析ツールのアルゴリズム等、原理・特徴を理解できる。
	プロテオーム解析のための解析ツールの利用法の知識	ツールのオプションの意味を理解し、目的タンパク質の同定のためのマススペクトルの検索ができる。		
	<ul style="list-style-type: none"> <li>バックグラウンド評価</li> <li>スペクトルの帰属</li> <li>Mascot</li> </ul>	バックグラウンドのノイズを除いて、目的タンパク質由来のピークだけを抽出できる。基本的なオプションを利用できる。	ツールのオプションの概要を理解し、実験条件を考慮した最適なオプションを設定できる。	ツールの各オプションを十分に理解し、実験条件を考慮した最適なオプションを設定できる。
解析結果の解釈	分子生物学の知識	解析結果を分子生物学の知識で解釈できる。		
	<ul style="list-style-type: none"> <li>分子生物学用語の知識 <ul style="list-style-type: none"> <li>タンパク質修飾</li> </ul> </li> <li>文献読解</li> <li>Gene ontology</li> </ul>	解析結果中の基本的な専門用語が理解できる。	解析結果中の専門用語が理解でき、内容を取捨選択し、総合的な判断ができる。関連する文献内容が理解できる。	解析結果中の専門用語が理解でき、内容を取捨選択し、総合的な判断ができる。関連する文献内容が理解できる。
	プロテオーム解析のための統計学の知識	解析結果の数値から結果の信頼性（統計的有意性）が判断できる。		
	<ul style="list-style-type: none"> <li>質量誤差の評価</li> <li>他の解析法との比較</li> </ul>	実測と計算との質量誤差から予測精度を評価できる。	解析結果の判断に必要な統計用語を理解できる。実測と計算との質量誤差から予測精度を評価できる。	解析結果の判断に必要な統計用語を理解できる。実測と計算との質量誤差から予測精度を評価できる。
	プロテオーム解析のための総合判断・産業応用の知識	解析結果を総合的に判断し、産業への応用を検討できる。		

	<ul style="list-style-type: none"> <li>機能推定</li> <li>他の解析法との比較</li> <li>医療及び産業応用に関する知識</li> </ul>		<p>検索・予測結果から配列の機能を推定できる。 オプション変更や他のツールでの再検索により、結果を比較検討できる。</p>	<p>検索・予測結果から配列の機能を推定できる。 オプション変更や他のツールで再検索を行い、結果を比較検討できる。 疾患関連遺伝子等、機能遺伝子の産業への応用を検討できる。</p>
--	---	--	--	--

4 2 5 タンパク質間ネットワーク(遺伝子ネットワーク)解析(大規模解析に適用されている実験方法は最終的に質量分析法を用いる)

業務内容	知識項目	スキル項目		
		初級	中級	上級
		(但し、初級・中級・上級の違いは理解すべき項目や理解すべきレベルによって差別化する。)		
実験の理解	タンパク質間ネットワーク解析の実験目的、実験原理の知識	実験目的、実験原理を理解できる。		
	<ul style="list-style-type: none"> <li>・ タンパク質間相互作用</li> <li>・ 代謝パスウェイ</li> <li>・ 酵母 two-hybrid 法                             <ul style="list-style-type: none"> <li>➢ cDNA ライブラリーからのスクリーニング</li> <li>➢ 転写調節因子</li> <li>➢ レポーター遺伝子</li> </ul> </li> <li>・ プルダウン法</li> <li>・ プロテインチップ</li> </ul>	実験目的、実験原理の概略を理解できる。	実験目的、実験原理を十分理解できる。	実験目的、実験原理を十分理解できる。
	タンパク質間ネットワーク解析の実験装置、実験方法の知識	実験装置、実験方法を理解できる。		
	<ul style="list-style-type: none"> <li>・ 酵母 two-hybrid 法</li> <li>・ プルダウン法                             <ul style="list-style-type: none"> <li>➢ 質量分析法</li> <li>➢ ウェスタンブロット法</li> </ul> </li> <li>・ プロテインチップ                             <ul style="list-style-type: none"> <li>➢ 質量分析法</li> </ul> </li> <li>・ プロテアーゼの基質特異性</li> </ul>	実験装置、実験方法の概略を理解できる。	実験装置、実験方法を理解できる。プロテアーゼの基質特異性、生体内で起こり得るタンパク質修飾等について理解できる。	実験装置を用いて対象データを得るプロセスを十分理解しており、データの信頼性について評価できる。MALDI-TOF、LC-MS 等、各実験方法の特徴を把握している。
解析ツールを用いたデータ解析	タンパク質間ネットワーク解析のための検索データベースの知識	最新かつ最大のアミノ酸配列データベースを把握しており、検索に利用できる。		

	<ul style="list-style-type: none"> <li>複数のデータベースの知識・比較</li> <li>データ量</li> <li>データの冗長性</li> <li>網羅度</li> <li>更新間隔</li> </ul>	指定されたデータベースの概要を理解できる。	代表的なデータベースの特徴を理解できる。	十分な種類のデータベースの特徴を理解しており、それらから最適なデータベースを選択できる。
	タンパク質間ネットワーク解析のための解析ツールの原理の知識	解析ツールのアルゴリズム等、原理・特徴を理解し、最適なツールを選択できる。		
	<ul style="list-style-type: none"> <li>PMF (peptide mass finger printing) 法</li> <li>MS/MS ( tandem MS ) 法</li> </ul>	解析ツールのアルゴリズム等、原理・特徴の概略を理解できる。	解析ツールのアルゴリズム等、原理・特徴の概略を理解できる。	解析ツールのアルゴリズム等、原理・特徴を理解できる。
	タンパク質間ネットワーク解析のための解析ツールの利用法の知識	ツールのオプションの意味を理解し、目的タンパク質の同定のためのマススペクトルの検索ができる。		
	<ul style="list-style-type: none"> <li>バックグラウンド評価</li> <li>スペクトルの帰属</li> <li>Mascot</li> </ul>	バックグラウンドのノイズを除いて、目的タンパク質由来のピークだけを抽出できる。 基本的なオプションを利用できる。	ツールのオプションの概要を理解し、実験条件を考慮した最適なオプションを設定できる。	ツールの各オプションを十分に理解し、実験条件を考慮した最適オプションを設定できる。
解析結果の解釈	分子生物学の知識	解析結果を分子生物学の知識で解釈できる。		

<ul style="list-style-type: none"> <li>分子生物学用語の知識 <ul style="list-style-type: none"> <li>タンパク質修飾</li> </ul> </li> <li>文献読解</li> <li>Gene ontology</li> <li>遺伝子ネットワーク</li> <li>システムバイオロジー</li> </ul>	<p>解析結果中の基本的な専門用語が理解できる。</p>	<p>解析結果中の専門用語が理解でき、内容を取捨選択し、総合的な判断ができる。 関連する文献内容が理解できる。</p>	<p>解析結果中の専門用語が理解でき、内容を取捨選択し、総合的な判断ができる。 関連する文献内容が理解できる。</p>
<p>タンパク質間ネットワーク解析のための統計学の知識</p>	<p>解析結果の数値から結果の信頼性（統計的有意性）が判断できる。</p>		
<ul style="list-style-type: none"> <li>質量誤差の評価</li> <li>他の解析法との比較</li> </ul>	<p>解析結果の判断に必要な統計用語の概略を理解できる。</p>	<p>解析結果の判断に必要な統計用語を理解できる。 実測と計算との質量誤差から予測精度を評価できる。</p>	<p>解析結果の判断に必要な統計用語を理解できる。 実測と計算との質量誤差から予測精度を評価できる。 オプション変更や他のツールで再検索を行い、結果を比較検討できる。</p>
<p>タンパク質間ネットワーク解析のための総合判断・産業応用の知識</p>	<p>解析結果を総合的に判断し、産業への応用を検討できる。</p>		
<ul style="list-style-type: none"> <li>機能推定</li> <li>検索・予測結果の比較・検討</li> <li>医療及び産業応用に関する知識</li> <li>遺伝子ネットワーク</li> <li>システムバイオロジー</li> </ul>		<p>結果から配列の機能を推定できる。 オプション変更や他のツールでの再検索により、結果を比較検討できる。</p>	<p>結果から配列の機能を推定できる。 オプション変更や他のツールでの再検索により、結果を比較検討できる。 相互作用やネットワーク解析も含めて、疾患関連遺伝子等、機能遺伝子の産業への応用を検討できる。</p>

4 2 6 資質・志向

業務内容	知識項目	スキル項目		
		初級	中級	上級
		(但し、初級・中級・上級の違いは理解すべき項目や理解すべきレベルによって差別化する。)		
資質・志向	データベースの国際的な基準への準拠	データベースの国際的な基準を理解し、作業に反映できる。		
	<ul style="list-style-type: none"> <li>データベースの最新情報</li> <li>国内外の研究機関との連携</li> <li>最新文献の理解</li> </ul>	代表的なデータベースの特徴を理解できる。	最新文献を理解しており、作業に反映できる。	最新文献を理解しており、作業に反映できる。 データベースの最新情報を把握し、データベースの作成時の立案、管理ができる。
	データの信頼度を高める一貫性、継続性、確実性	統一ルールを理解し、それに従って作業できる。		
	<ul style="list-style-type: none"> <li>統一ルールの遵守</li> <li>統一ルールの作成</li> </ul>	統一ルールを理解し、それに従って作業できる。	グループ内での作業の統一性を維持し、統括できる。	統一ルールを作成できる。
	生命科学の急速な進展と並走する前進性	最新文献を検索かつ理解し、作業に反映できる。		
	<ul style="list-style-type: none"> <li>最新文献の理解</li> </ul>	最新文献を検索できる。	最新文献を理解しており、作業に反映できる。	最新文献を理解しており、それに基づいて、入力された記述の正誤について検討できる。
	実務の倫理性	個人情報保護・生命倫理・工業技術者倫理・薬事制度・ISO に関する規制の存在を理解し、作業の中で遵守できる。		

	<ul style="list-style-type: none"> <li>・ 個人情報の保護</li> <li>・ 遺伝子取り扱い倫理規定</li> </ul>	<p>公開情報と非公開情報を区別できる。</p>	<p>公開情報と非公開情報を区別できる。</p>	<p>公開情報と非公開情報を区別できる。 倫理的なルールを作成、指導できる。</p>
--	--	--------------------------	--------------------------	--

## 第5章 カリキュラム

### 5 1 . 教育コース

前章で策定したスキルスタンダードのレベル分けに合わせ、以下の通り教育コースを初級・中級・上級に分類し、設定した。濃色であるほど、高度な内容であることを示す。

このレベル設定では、すべての実験において「解析結果の解釈」の中の「総合判断・産業応用」は中級および上級だけができるものとした。さらに、中級と上級間ではレベル差を設けて、上級だけが解析結果からの応用利用を考えることができるものとした。

図表 5-1 教育コースとレベル設定

実験項目	科目名	初級	中級	上級
分子生物学	分子生物学			
配列解析	配列解析実験			
	配列解析におけるデータ解析			
	配列解析におけるデータ解釈 (統計学)			
	配列解析におけるデータ解釈 (総合判断・産業応用)			
発現プロファイル解析	発現プロファイル解析実験			
	発現プロファイル解析におけるデータ解析			
	発現プロファイル解析におけるデータ解釈 (統計学)			
	発現プロファイル解析におけるデータ解釈 (総合判断・産業応用)			
P) 多型解析(SN)	多型解析実験			
	多型解析におけるデータ解析			
	多型解析におけるデータ解釈 (統計学)			
	多型解析におけるデータ解釈 (総合判断・産業応用)			
解析 プロテオーム	プロテオーム解析実験			
	プロテオーム解析におけるデータ解析			
	プロテオーム解析におけるデータ解釈 (統計学)			
	プロテオーム解析におけるデータ解釈 (総合判断・産業応用)			
解析 タンパク質間ネットワーク	タンパク質間ネットワーク解析実験			
	タンパク質間ネットワーク解析におけるデータ解析			
	タンパク質間ネットワーク解析におけるデータ解釈 (統計学)			
	タンパク質間ネットワーク解析におけるデータ解釈 (総合判断・産業応用)			
資質・志向	アノテーター序説			

## 5 2 .カリキュラム

### 5 2 1 アノテーション・スタンダード（アノテーター初級）人材育成コース

#### （1）育成のねらい

大学生物系学部卒の人材がコンピュータースキルの基礎を学んだ上で、データ解析の方法を習得し、リーダーからの指示通りにアノテーション作業ができる人材を育成することをねらいとする。

#### （2）育成目標（アノテーション・スタンダードの役割・能力・知識）

- ・アノテーションチームの初級者として、リーダーからの指示を的確に理解してアノテーション作業を行うことができる。
- ・特定の実験分野に限って最低限必要とされる知識を持ち、他の研究者からの指示を理解し、実行できる。
- ・大量データ処理の過程で一連の作業を同じ基準、同じ様式で繰り返し処理できる。

#### （3）履修科目

図表 5-2 アノテーション・スタンダード人材育成コース履修科目

実験区分	科目名	内容
分子生物学	分子生物学基礎	データ解釈のために必要な分子生物学の基礎、文献読解、Gene ontologyの利用について習得する。
配列解析	配列解析実験概論	配列解析における実験目的、実験原理、実験装置、実験方法の概要を習得する。
	配列解析におけるデータ解析	配列解析に必要な検索データベースの選択、検索・予測ツールの選択のための知識を習得し、ツールを使いこなす方法を習得する。
	配列解析におけるデータ解釈（統計学基礎）	配列解析において解析結果の解釈に必要な統計学の基礎を習得する。
発現プロファイル解析	発現プロファイル解析実験概論	発現プロファイル解析における実験目的、実験原理、実験装置、実験方法の概要を習得する。
	発現プロファイル解析におけるデータ解析	発現プロファイル解析において解析ツールの選択に必要な知識を習得し、ツールを使いこなす方法を習得する。
	発現プロファイル解析におけるデータ解釈（統計学基礎）	発現プロファイル解析において解析結果の解釈に必要な統計学の基礎を習得する。
多型解析（SNP）	多型解析実験概論	多型解析における実験目的、実験原理、実験装置、実験方法の概要を習得する。

	多型解析におけるデータ解析	多型解析に必要な検索データベースの選択、検索・予測ツールの選択のための知識を習得し、ツールを使いこなす方法を習得する。
	多型解析におけるデータ解釈（統計学基礎）	多型解析において解析結果の解釈に必要な統計学の基礎を習得する。
プロテオーム解析	プロテオーム解析実験概論	プロテオーム解析における実験目的、実験原理、実験装置、実験方法の概要を習得する。
	プロテオーム解析におけるデータ解析	プロテオーム解析において解析ツールの選択に必要な知識を習得し、ツールを使いこなす方法を習得する。
	プロテオーム解析におけるデータ解釈（統計学基礎）	プロテオーム解析において解析結果の解釈に必要な統計学の基礎を習得する。
タンパク質間ネットワーク解析	タンパク質間ネットワーク解析実験概論	タンパク質間ネットワーク解析における実験目的、実験原理、実験装置、実験方法の概要を習得する。
	タンパク質間ネットワーク解析におけるデータ解析	タンパク質間ネットワーク解析において解析ツールの選択に必要な知識を習得し、ツールを使いこなす方法を習得する。
	タンパク質間ネットワーク解析におけるデータ解釈（統計学基礎）	タンパク質間ネットワーク解析において解析結果の解釈に必要な統計学の基礎を習得する。
資質・志向	アノテーター序説	データベースの国際的な基準への準拠、データの信頼度を高める一貫性・継続性・確実性、生命科学の急速な進展と並走する前進性、実務の倫理性について習得する。

## 5 2 2 アノテーション・エキスパート（アノテーター中級）人材育成コース

### （1）育成のねらい

遺伝子やタンパク質のゲノムワイドな解析実験から生み出される大量の生物データを解析し、各データにアノテーションを行う過程において、アノテーションチームの中核として、リーダーからの指導なしに自立してアノテーションを行うことができる人材を育成することをねらいとする。

### （2）育成目標（アノテーション・エキスパートの役割・能力・知識）

- ・アノテーションチームの中核として、自立してアノテーションを行うことができる。
- ・特定の専門分野に明るく、初級レベルと上級レベルをつなぐコミュニケーション能力を持ち、チームの中で橋渡し役の役割を果たす。
- ・実験のプロセス、複数の解析手法（解析ツール）を把握しており、解析結果を多面的に評価できる。

### （3）履修科目

図表 5-3 アノテーション・エキスパート人材育成コース履修科目

実験項目	科目名	内容
分子生物学	分子生物学概論	データ解釈のために必要な分子生物学の概要、文献読解、Gene ontologyの利用について習得する。
配列解析	配列解析実験特論	配列解析における実験目的、実験原理、実験装置、実験方法を習得する。特に生データの精度まで判断できる方法を習得する。
	配列解析におけるデータ解析	配列解析に必要な検索データベースの選択、検索・予測ツールの選択のための知識を習得し、ツールを使いこなす方法を習得する。
	配列解析におけるデータ解釈（統計学）	配列解析において解析結果の解釈に必要な統計学を習得する。
	配列解析におけるデータ解釈（総合判断）	配列解析において解析結果の解釈に必要な機能推定、複数の解析方法による結果の比較・検討について習得する。
発現プロファイル解析	発現プロファイル解析実験特論	発現プロファイル解析における実験目的、実験原理、実験装置、実験方法を習得する。特に生データの精度まで判断できる方法を習得する。
	発現プロファイル解析におけるデータ解析	発現プロファイル解析において解析ツールの選択に必要な知識を習得し、ツールを使いこなす方法を習得する。
	発現プロファイル解析におけるデータ解釈（統計学）	発現プロファイル解析において解析結果の解釈に必要な統計学を習得する。

	発現プロファイル解析におけるデータ解釈（総合判断）	発現プロファイル解析において解析結果の解釈に必要な機能推定、複数の解析方法による結果の比較・検討について習得する。
多型解析（SNP）	多型解析実験特論	多型解析における実験目的、実験原理、実験装置、実験方法を習得する。特に生データの精度まで判断できる方法を習得する。
	多型解析におけるデータ解析	多型解析に必要な検索データベースの選択、検索・予測ツールの選択のための知識を習得し、ツールを使いこなす方法を習得する。
	多型解析におけるデータ解釈（統計学）	多型解析において解析結果の解釈に必要な統計学を習得する。
	多型解析におけるデータ解釈（総合判断）	多型解析において解析結果の解釈に必要な機能推定、複数の解析方法による結果の比較・検討について習得する。
プロテオーム解析	プロテオーム解析実験特論	プロテオーム解析における実験目的、実験原理、実験装置、実験方法を習得する。特に生データの精度まで判断できる方法を習得する。
	プロテオーム解析におけるデータ解析	プロテオーム解析において解析ツールの選択に必要な知識を習得し、ツールを使いこなす方法を習得する。
	プロテオーム解析におけるデータ解釈（統計学）	プロテオーム解析において解析結果の解釈に必要な統計学を習得する。
	プロテオーム解析におけるデータ解釈（総合判断）	プロテオーム解析において解析結果の解釈に必要な機能推定、複数の解析方法による結果の比較・検討について習得する。
タンパク質間ネットワーク解析	タンパク質間ネットワーク解析実験特論	タンパク質間ネットワーク解析における実験目的、実験原理、実験装置、実験方法を習得する。特に生データの精度まで判断できる方法を習得する。
	タンパク質間ネットワーク解析におけるデータ解析	タンパク質間ネットワーク解析において解析ツールの選択に必要な知識を習得し、ツールを使いこなす方法を習得する。
	タンパク質間ネットワーク解析におけるデータ解釈（統計学）	タンパク質間ネットワーク解析において解析結果の解釈に必要な統計学を習得する。

	タンパク質間ネットワーク解析におけるデータ解釈（総合判断）	タンパク質間ネットワーク解析において解析結果の解釈に必要な機能推定、複数の解析方法による結果の比較・検討について習得する。
資質・志向	アノテーター序説	データベースの国際的な基準への準拠、データの信頼度を高める一貫性・継続性・确实性、生命科学の急速な進展と並走する前進性、実務の倫理性について習得する。

## 5 2 3 アノテーション・マネージャー（アノテーター上級）人材育成コース

### （1）育成のねらい

遺伝子やタンパク質のゲノムワイドな解析実験から生み出される大量の生物データを解析し、各データにアノテーションを行う過程において、アノテーションチームのリーダーとして、総合的な視点からプロジェクトの立案、推進、管理まで行うことができる人材を育成することをねらいとする。

### （2）育成目標（アノテーション・マネージャーの役割・能力・知識）

- ・アノテーションチームのリーダーとして、総合的な視点からプロジェクトの立案、推進、管理を行うことができる。
- ・特定の専門分野に限らず幅広い知識とコミュニケーション能力を持ち、研究者（実験研究者、バイオインフォマティクス研究者）やデータベースシステム開発者などの共同構成員の中で橋渡し役として中心的な役割を果たす。

### （3）履修科目

図表 5-4 アノテーション・マネージャー人材育成コース履修科目

実験項目	科目名	内容
分子生物学	分子生物学概論	データ解釈のために必要な分子生物学の概要、文献読解、Gene ontologyの利用について習得する。
配列解析	配列解析実験特論	配列解析における実験目的、実験原理、実験装置、実験方法を習得する。特に生データの精度まで判断できる方法を習得する。
	配列解析におけるデータ解析	配列解析に必要な検索データベースの選択、検索・予測ツールの選択のための知識を習得し、ツールを使いこなす方法を習得する。
	配列解析におけるデータ解釈（統計学）	配列解析において解析結果の解釈に必要な統計学を習得する。
	配列解析におけるデータ解釈特論（総合判断・産業応用）	配列解析において解析結果の解釈に必要な機能推定、複数の解析方法による結果の比較・検討、機能の応用利用について習得する。
発現プロファイル解析	発現プロファイル解析実験特論	発現プロファイル解析における実験目的、実験原理、実験装置、実験方法を習得する。特に生データの精度まで判断できる方法を習得する。
	発現プロファイル解析におけるデータ解析	発現プロファイル解析において解析ツールの選択に必要な知識を習得し、ツールを使いこなす方法を習得する。
	発現プロファイル解析におけるデータ解釈（統計学）	発現プロファイル解析において解析結果の解釈に必要な統計学を習得する。

	発現プロファイル解析におけるデータ解釈特論（総合判断・産業応用）	発現プロファイル解析において解析結果の解釈に必要な機能推定、複数の解析方法による結果の比較・検討、機能の応用利用について習得する。
多型解析（SNP）	多型解析実験特論	多型解析における実験目的、実験原理、実験装置、実験方法を習得する。特に生データの精度まで判断できる方法を習得する。
	多型解析におけるデータ解析	多型解析に必要な検索データベースの選択、検索・予測ツールの選択のための知識を習得し、ツールを使いこなす方法を習得する。
	多型解析におけるデータ解釈（統計学）	多型解析において解析結果の解釈に必要な統計学を習得する。
	多型解析におけるデータ解釈特論（総合判断・産業応用）	多型解析において解析結果の解釈に必要な機能推定、複数の解析方法による結果の比較・検討、機能の応用利用について習得する。
プロテオーム解析	プロテオーム解析実験特論	プロテオーム解析における実験目的、実験原理、実験装置、実験方法を習得する。特に生データの精度まで判断できる方法を習得する。
	プロテオーム解析におけるデータ解析	プロテオーム解析において解析ツールの選択に必要な知識を習得し、ツールを使いこなす方法を習得する。
	プロテオーム解析におけるデータ解釈（統計学）	プロテオーム解析において解析結果の解釈に必要な統計学を習得する。
	プロテオーム解析におけるデータ解釈特論（総合判断・産業応用）	プロテオーム解析において解析結果の解釈に必要な機能推定、複数の解析方法による結果の比較・検討、機能の応用利用について習得する。
タンパク質間ネットワーク解析	タンパク質間ネットワーク解析実験特論	タンパク質間ネットワーク解析における実験目的、実験原理、実験装置、実験方法を習得する。特に生データの精度まで判断できる方法を習得する。
	タンパク質間ネットワーク解析におけるデータ解析	タンパク質間ネットワーク解析において解析ツールの選択に必要な知識を習得し、ツールを使いこなす方法を習得する。
	タンパク質間ネットワーク解析におけるデータ解釈（統計学）	タンパク質間ネットワーク解析において解析結果の解釈に必要な統計学を習得する。

	タンパク質間ネットワーク解析におけるデータ解釈特論（総合判断・産業応用）	タンパク質間ネットワーク解析において解析結果の解釈に必要な機能推定、複数の解析方法による結果の比較・検討、機能の応用利用について習得する。
資質・志向	アノテーター序説	データベースの国際的な基準への準拠、データの信頼度を高める一貫性・継続性・确实性、生命科学の急速な進展と並走する前進性、実務の倫理性について習得する。

### 5 3 . シラバス

<b>科目名称</b>	分子生物学基礎
<b>区分</b>	必修
<b>講義形態</b>	座学
<b>時間</b>	3 時間 ( 1 コマ )
<b>対象者</b>	カリキュラム初級コース受講者 Bio 系分野 / IT 系分野所属 専門学校または大学学部卒程度で分子生物学・生命科学を学んだ者
<b>対象知識項目</b>	分子生物学の知識
<b>達成目標</b>	データ解釈のために必要な分子生物学の基礎、文献読解、Gene ontology の利用について習得する。
<b>科目概要</b>	1 分子生物学用語の知識 2 文献読解 3 Gene ontology
<b>参考文献・資料</b>	ゲノム 2 新しい生命情報システムへのアプローチ T.A. Brown (著), 村松 正実 (翻訳) メディカルサイエンスインターナショナル (2003)  細胞の分子生物学 第 4 版 Bruce Alberts (著), 中村 桂子 (翻訳), 松原 謙一 (翻訳) ニュートンプレス (2004)
<b>達成度評価の方法</b>	テストによる評価
<b>備考</b>	

<b>科目名称</b>	配列解析実験概論
<b>区分</b>	必修
<b>講義形態</b>	座学
<b>時間</b>	3 時間 ( 1 コマ )
<b>対象者</b>	カリキュラム初級コース受講者 Bio 系分野 / IT 系分野所属 専門学校または大学学部卒程度で分子生物学・生命科学を学んだ者
<b>対象知識項目</b>	配列解析の実験目的、実験原理の知識、配列解析の実験装置、実験方法の知識
<b>達成目標</b>	配列解析における実験目的、実験原理、実験装置、実験方法の概要を習得する。
<b>科目概要</b>	1 基礎的な背景知識 1.1 クローニングの方法 1.2 シーケンスの原理と方法  2 実験の原理・方法 2.1 Base call の原理と方法 2.2 Assembly の原理と方法 2.3 ショットガンシーケンス
<b>参考文献・資料</b>	これからのバイオインフォマティクスのためのバイオ実験入門 実験の原理から理解するバイオの基礎 大藤 道衛 (編集), 高井 貴子 (編集), 高木 利久 羊土社 (2002)
<b>達成度評価の方法</b>	テストによる評価
<b>備考</b>	

<b>科目名称</b>	配列解析におけるデータ解析
-------------	---------------

区分	必修
講義形態	座学と実習
時間	3時間(1コマ)
対象者	カリキュラム初級/中級/上級コース受講者 Bio系分野/IT系分野所属 専門学校または大学学部卒以上で分子生物学・生命科学を学んだ者
対象知識項目	配列解析のための検索データベースの知識、配列解析のための検索・予測ツールの原理の知識、配列解析のための検索・予測ツールの利用法の知識
達成目標	配列解析に必要な検索データベースの選択、検索・予測ツールの選択のための知識を習得し、ツールを使いこなす方法を習得する。
科目概要	<p>1 配列編集</p> <p>1.1 Base caller による配列決定</p> <p>1.1.1 Phred</p> <p>1.2 Assembly</p> <p>1.2.1 ベクター・アダプター配列の除去</p> <p>1.2.2 Phrap</p> <p>2 解析方法(解析ツール)の原理</p> <p>2.1 ホモロジー検索</p> <p>2.1.1 配列データベースの知識</p> <p>2.1.1.1 代表的な配列データベースとその特徴</p> <p>2.1.1.2 検索対象データベースの選択</p> <p>2.1.1.2.1 検索時のチェック事項(データ量、データの冗長性、網羅度、更新間隔)</p> <p>2.1.2 検索プログラムの知識(BLAST, FASTA, SSEARCH)</p> <p>2.1.2.1 オプション</p> <p>2.1.2.2 アルゴリズム</p> <p>2.1.3 フィルター</p> <p>2.1.3.1 Repeat Masker</p> <p>2.2 遺伝子予測</p> <p>2.2.1 遺伝子構造</p> <p>2.2.2 各ツールの特徴(Genescan, Grail, GeneMark, Glimmer, GlimmerHMM, FGENES, Spidy, ALN)</p> <p>2.2.2.1 対象とする生物種(真核生物、原核生物)</p> <p>2.2.2.2 サービスを提供しているサイト</p> <p>2.2.3 アルゴリズム(ab initio 法、EST/cDNA とゲノム配列とのアライメント、隠れマルコフモデル(HMM))</p> <p>2.2.4 比較ゲノム解析</p> <p>2.3 モチーフ検索</p> <p>2.4 プロファイル(ドメイン)検索</p> <p>2.4.1 細胞内局在予測</p> <p>2.4.1.1 シグナルペプチド</p> <p>2.4.2 膜貫通領域</p> <p>2.5 二次構造予測</p> <p>2.6 立体構造予測(スレッディング)</p> <p>3 解析方法(解析ツール)の実習</p> <p>3.1 BLAST(ホモロジー検索)</p> <p>3.1.1 BLASTの代表的なバージョン(NCBI/BLAST、WU-BLAST)</p> <p>3.1.2 BLASTファミリーの選択(BLASTN, BLASTP, TBLASTN, BLASTX, TBLASTX)</p> <p>3.1.3 オプション(パラメーター/しきい値)の意味、設定(使い分け)</p> <p>3.1.4 検索対象データベース、検索対象 subdivision(nr、生物種、EST等)</p> <p>3.1.5 出力形式、結果一覧表示数、アラインメント表示数</p>

	<ul style="list-style-type: none"> <li>3.1.6 期待値 (E-value、p-value) フィルター (SEG、DUST) ワードサイズ</li> <li>3.1.7 アミノ酸置換テーブル (BLOSUM、PAM) 開始ギャップペナルティー、伸張ギャップペナルティー</li> <li>3.1.8 BLAST/FASTA パーザー (結果の構文解析プログラム)</li> <li>3.1.9 Bioperl, Bioruby</li> <li>3.2 Genescan (遺伝子予測) <ul style="list-style-type: none"> <li>3.2.1 オプション</li> <li>3.2.2 生物種 (真核生物、原核生物)</li> <li>3.2.3 Suboptimal exon cutoff</li> <li>3.2.4 Initial exon, Internal exon, Terminal exon</li> <li>3.2.5 Single-exon gene</li> </ul> </li> <li>3.3 予測精度</li> <li>3.4 具体的なツール (比較ゲノム) <ul style="list-style-type: none"> <li>3.4.1 megablast</li> <li>3.4.2 Artemis Comparison Tool</li> </ul> </li> </ul>
<p><b>参考文献・資料</b></p>	<p>生命情報学キートン  D.R. Westhead, J.H. Parish, R.M. Twyman  監訳: 池尾一穂, 鈴木善幸, 五條堀 孝  シュプリンガーフェアラーク東京 (2003)</p> <p>初心者でもわかる!バイオインフォマティクス入門 やさしい UNIX 操作から遺伝子・タンパク質解析まで  坊農 秀雅 羊土社 (2002)</p> <p>バイオインフォマティクス 応用生命科学シリーズ9  美宅成樹, 榊 佳之 東京化学同人 (2003)</p> <p>実践 バイオインフォマティクス -ゲノム研究のためのコンピュータスキル-  Cynthia Gibas (著), Per Jambeck (著), 水島 洋 (翻訳), 明石 浩史 (翻訳),  またぬき (翻訳)  オライリー・ジャパン (2002)</p> <p>バイオインフォマティクス ゲノム配列から機能解析へ  岡崎 康司 (翻訳), 坊農 秀雅 (翻訳), David W. Mount  メディカル・サイエンス・インターナショナル (2002)</p> <p>バイオデータベースとウェブツールの手とり足とり活用法  中村保一, 磯合 敦, 石川 淳 羊土社 (2003)</p> <p>ゲノムネットのデータベース利用法 第3版  金久 実 (編) 共立出版 (2002)</p> <p>即活用のためのバイオインフォマティクス入門  美宅 成樹, 広川 貴次 中山書店 (2004)</p> <p>Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins  (Methods of Biochemical Analysis, V. 43)  Edited by: Andreas D. Baxevanis and B. F. Francis Ouellette  WILEY-INTERSCIENCE (2001)</p> <p>BLAST: at the core of a powerful and diverse set of sequence analysis tools.  Scott McGinnis and Thomas L. Madden  Nucleic Acids Research 2004, 32(Web Server issue):W20-W25</p>

	分子進化 解析の技法とその応用 宮田隆 共立出版 (1998)
	生命情報学 五條堀隆 シュプリンガーフェアラーク東京 (2003)
	分子進化遺伝学 根井正利 (五條堀隆、斎藤成也翻訳) 培風館 (1990)
達成度評価の方法	テストによる評価
備考	

科目名称	配列解析におけるデータ解釈 (統計学基礎)
区分	必修
講義形態	座学と実習
時間	3 時間 (1 コマ)
対象者	カリキュラム初級コース受講者 Bio 系分野 / IT 系分野所属 専門学校または大学学部卒程度で分子生物学・生命科学を学んだ者
対象知識項目	配列解析のための統計学の知識
達成目標	配列解析において解析結果の解釈に必要な統計学の基礎を習得する。
科目概要	1 解析結果の解釈 1.1 統計的有意性 1.1.1 確率分布 1.1.1.1 ポワソン分布、E-value、p-value 1.1.2 標準偏差 1.1.2.1 Z score 1.1.3 尤度
参考文献・資料	バイオインフォマティクス ゲノム配列から機能解析へ 岡崎 康司 (翻訳), 坊農 秀雅 (翻訳), David W. Mount メディカル・サイエンス・インターナショナル (2002)
	Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins (Methods of Biochemical Analysis, V. 43) Edited by: Andreas D. Baxevanis & B. F. Francis Ouellette WILEY-INTERSCIENCE (2001)
	BLAST: at the core of a powerful and diverse set of sequence analysis tools. Scott McGinnis and Thomas L. Madden Nucleic Acids Research 2004, 32(Web Server issue):W20-W25
達成度評価の方法	テストによる評価
備考	

科目名称	発現プロファイル解析実験概論
区分	必修
講義形態	座学
時間	3 時間 (1 コマ)
対象者	カリキュラム初級コース受講者 Bio 系分野 / IT 系分野所属 専門学校または大学学部卒程度で分子生物学・生命科学を学んだ者
対象知識項目	発現プロファイル解析の実験目的、実験原理の知識、発現プロファイル解析の実験装置、実験方法の知識

<b>達成目標</b>	発現プロファイル解析における実験目的、実験原理、実験装置、実験方法の概要を習得する。
<b>科目概要</b>	<ol style="list-style-type: none"> <li>1 基礎的な背景知識 <ol style="list-style-type: none"> <li>1.1 遺伝子発現のしくみ</li> <li>1.2 ハイブリダイゼーションの原理</li> <li>1.3 蛍光ラベリングの原理</li> </ol> </li> <li>2 実験の原理・方法 <ol style="list-style-type: none"> <li>2. DNAマイクロアレイ、DNAチップの種類 <ol style="list-style-type: none"> <li>2.1. チップ型アレイ (GeneChip)</li> <li>2.2. スポット型アレイ (cDNA 型アレイ)</li> </ol> </li> </ol> </li> <li>3 実験準備の知識 <ol style="list-style-type: none"> <li>3.1 調整試料 <ol style="list-style-type: none"> <li>3.1.1 cDNA ライブラリー</li> <li>3.1.2 RNA 抽出</li> <li>3.1.3 PCR</li> </ol> </li> <li>3.2 試薬 <ol style="list-style-type: none"> <li>3.2.1 蛍光色素 (Cy3, Cy5)</li> </ol> </li> <li>3.3 機材 <ol style="list-style-type: none"> <li>3.3.1 スキャナー</li> <li>3.3.2 アレイ / チップ上の遺伝子の既知情報</li> </ol> </li> </ol> </li> </ol>
<b>参考文献・資料</b>	<p>これからのバイオインフォマティクスのためのバイオ実験入門 実験の原理から理解するバイオの基礎 大藤 道衛 (編集), 高井 貴子 (編集), 高木 利久 羊土社 (2002)</p> <p>DNA マイクロアレイ実践マニュアル 林崎良英 (監修), 岡崎康司 (編) 羊土社 (2000)</p>
<b>達成度評価の方法</b>	テストによる評価
<b>備考</b>	

<b>科目名称</b>	発現プロファイル解析におけるデータ解析
<b>区分</b>	必修
<b>講義形態</b>	座学と実習
<b>時間</b>	3 時間 ( 1 コマ )
<b>対象者</b>	カリキュラム初級 / 中級 / 上級コース受講者 Bio 系分野 / IT 系分野所属 専門学校または大学学部卒以上で分子生物学・生命科学を学んだ者
<b>対象知識項目</b>	発現プロファイル解析のための解析ツールの知識、発現プロファイル解析のための解析ツールの原理の知識、発現プロファイル解析のための解析ツールの利用法の知識
<b>達成目標</b>	発現プロファイル解析において解析ツールの選択に必要な知識を習得し、ツールを使いこなす方法を習得する。
<b>科目概要</b>	<ol style="list-style-type: none"> <li>1 解析方法 ( 解析ツール ) の原理 <ol style="list-style-type: none"> <li>1.1 蛍光強度の定量化 <ol style="list-style-type: none"> <li>1.1.1 正規化</li> <li>1.1.2 バックグラウンド評価</li> <li>1.1.3 はずれ値検出</li> </ol> </li> <li>1.2 検定 <ol style="list-style-type: none"> <li>1.2.1 発現量変化の計算</li> <li>1.2.2 発現量変化の有意性の評価</li> </ol> </li> <li>1.3 クラスタ解析 <ol style="list-style-type: none"> <li>1.3.1 クラスタ化法の種類</li> </ol> </li> </ol> </li> </ol>

	<p>1.3.1.1 階層型クラスター化法、K-means クラスター化法、自己組織化マップ (SOM)</p> <p>1.3.2 相関解析</p> <p>1.3.3 多変量解析</p> <p>2 解析方法 (解析ツール) の実習</p> <p>2.1 データフォーマット</p> <p>2.1.1 CSV 形式</p> <p>2.1.2 MAGE-ML</p> <p>2.2 蛍光強度の数値化</p> <p>2.2.1 ScanAlyze</p> <p>2.3 PRIM</p> <p>2.4 クラスター分析・視覚化表示</p> <p>2.4.1 Cluster</p> <p>2.4.2 TreeView</p> <p>2.5 多変量解析ソフト</p> <p>2.5.1 R</p> <p>2.5.2 GeneSpring</p> <p>2.6 機能アノテーション</p> <p>2.6.1 NetAffx</p> <p>2.7 データマイニング</p> <p>2.7.1 Spotfire</p>
参考文献・資料	<p>初心者でもわかる!バイオインフォマティクス入門 やさしい UNIX 操作から遺伝子・タンパク質解析まで 坊農 秀雅 羊土社 (2002)</p> <p>バイオデータベースとウェブツールの手とり足とり活用法 中村保一, 磯合 敦, 石川 淳 羊土社 (2003)</p> <p>統合ゲノミクスのためのマイクロアレイデータアナリシス Kohane, I.S., Kho, A.T., Butte, A.J. (著), 星田有人 (訳) シュプリンガーフェアラーク東京 (2004)</p> <p>DNA マイクロアレイデータ解析入門 Knudsen, S. (著), 塩島聡, 松本治, 辻本豪三 (監訳) 羊土社 (2002)</p> <p>DNA マイクロアレイ実践マニュアル 林崎良英 (監修), 岡崎康司 (編) 羊土社 (2000)</p>
達成度評価の方法	テストによる評価
備考	

科目名称	発現プロファイル解析におけるデータ解釈 (統計学基礎)
区分	必修
講義形態	座学と実習
時間	3 時間 (1 コマ)
対象者	カリキュラム初級コース受講者 Bio 系分野 / IT 系分野所属 専門学校または大学学部卒程度で分子生物学・生命科学を学んだ者
対象知識項目	発現プロファイル解析のための統計学の知識
達成目標	発現プロファイル解析において解析結果の解釈に必要な統計学の基礎を習得する。
科目概要	1 解析結果の解釈 1.1 クラスター解析

	<ul style="list-style-type: none"> <li>1.1.1 クラスタ化法の種類 <ul style="list-style-type: none"> <li>1.1.1.1 階層型クラスタ化法、K-means クラスタ化法、自己組織化マップ (SOM)</li> </ul> </li> <li>1.1.2 相関解析</li> <li>1.1.3 多変量解析</li> </ul> <ul style="list-style-type: none"> <li>2.1 多変量解析ソフト <ul style="list-style-type: none"> <li>2.1.1 R</li> <li>2.1.2 GeneSpring</li> </ul> </li> </ul>
参考文献・資料	Rによる統計解析の基礎 Computer in Education and Research 7 中澤 港 ピアソンエデュケーション (2003)
達成度評価の方法	テストによる評価
備考	

科目名称	多型解析実験概論
区分	必修
講義形態	座学
時間	3時間 (1コマ)
対象者	カリキュラム初級コース受講者 Bio系分野/IT系分野所属 専門学校または大学学部卒程度で分子生物学・生命科学を学んだ者
対象知識項目	多型解析の実験目的、実験原理の知識、多型解析の実験装置、実験方法の知識
達成目標	多型解析における実験目的、実験原理、実験装置、実験方法の概要を習得する。
科目概要	<ul style="list-style-type: none"> <li>1 基礎的な背景知識 <ul style="list-style-type: none"> <li>1.1 遺伝子多型の種類 <ul style="list-style-type: none"> <li>1.1.1 RFLP</li> <li>1.1.2 VNTR</li> <li>1.1.3 マイクロサテライト</li> <li>1.1.4 SNP</li> </ul> </li> <li>1.2 SNPの種類</li> <li>1.3 SNPと疾患 <ul style="list-style-type: none"> <li>1.3.1 単一遺伝子疾患</li> <li>1.3.2 多因子疾患</li> </ul> </li> <li>1.4 SNPと薬物応答性</li> </ul> </li> <li>2 実験の原理・方法 <ul style="list-style-type: none"> <li>2.1 SNPの同定法 <ul style="list-style-type: none"> <li>2.1.1 シーケンス</li> <li>2.1.2 Assembly</li> <li>2.1.3 Invader法</li> <li>2.1.4 ICAM法</li> <li>2.1.5 DNAマイクロアレイ、DNAチップ</li> </ul> </li> </ul> </li> </ul>
参考文献・資料	これからのバイオインフォマティクスのためのバイオ実験入門 実験の原理から理解するバイオの基礎 大藤 道衛 (編集), 高井 貴子 (編集), 高木 利久 羊土社 (2002)
達成度評価の方法	テストによる評価
備考	

科目名称	多型解析におけるデータ解析
区分	必修
講義形態	座学と実習

<b>時間</b>	3 時間 ( 1 コマ )
<b>対象者</b>	カリキュラム初級 / 中級 / 上級コース受講者 Bio 系分野 / IT 系分野所属 専門学校または大学学部卒以上で分子生物学・生命科学を学んだ者
<b>対象知識項目</b>	多型解析のための検索データベースの知識、多型解析のための検索・予測ツールの原理の知識、多型解析のための検索・予測ツールの利用法の知識
<b>達成目標</b>	多型解析に必要な検索データベースの選択、検索・予測ツールの選択のための知識を習得し、ツールを使いこなす方法を習得する。
<b>科目概要</b>	<p>1 解析方法 ( 解析ツール ) の原理</p> <p>1.1 複数のデータベースの特徴・利用方法・比較</p> <p>1.1.1 dbSNP</p> <p>1.1.2 JSNP</p> <p>1.1.3 HGVBase</p> <p>1.2 Base caller プログラム ( 配列精度の評価 )</p> <p>1.2.1 Phred</p> <p>1.3 Assemble プログラム</p> <p>1.3.1 Phrap</p> <p>1.4 マルチプルアライメント</p> <p>1.5 ホモロジー検索</p> <p>1.5.1 BLAT ( 一致配列検索 )</p> <p>1.6 遺伝子予測</p> <p>1.6.1 SNP の位置と遺伝子構造</p> <p>2 解析方法 ( 解析ツール ) の実習</p> <p>2.1 Base caller プログラム ( 配列精度の評価 )</p> <p>2.1.1 Phred</p> <p>2.2 Assemble プログラム</p> <p>2.2.1 Phrap</p> <p>2.3 マルチプルアライメント</p>
<b>参考文献・資料</b>	<p>初心者でもわかる!バイオインフォマティクス入門 やさしい UNIX 操作から遺伝子・タンパク質解析まで 坊農 秀雅 羊土社 ( 2002 )</p> <p>即活用のためのバイオインフォマティクス入門 美宅 成樹, 広川 貴次 中山書店 ( 2004 )</p> <p>ヒト遺伝子アノテーションの統合データベース ゲノム医学 Vol.4, No.4 山崎千里、今西規 メディカルレビュー社 ( 2004 )</p> <p>遺伝子情報統合データベース H-Invitational Database 藤井康之、今西規、五條堀孝 バイオ高性能機器新技術利用マニュアル、蛋白質核酸酵素増刊号 小原収、谷口寿章、市川哲生、猪飼篤、編 ( 2004 )</p> <p>Integrative Annotation of 21,037 Human Genes Validated by Full-Length cDNA Clones. Imanishi T, Itoh T, Suzuki Y, O'Donovan C, Fukuchi S, Koyanagi KO, Barrero RA, Tamura T, Yamaguchi-Kabata Y, Tanino M, et al. PLoS Biology 2: 856-875. ( 2004 )</p> <p>The Human Anatomic Gene Expression Library (H-ANGEL), the H-Inv Integrative Display of Human Gene Expression across Disparate Technologies and Platforms. Tanino M, Debily MA, Tamura T, Hishiki T, Ogasawara O, Murakawa K, Kawamoto S, Itoh K, Watanabe S, de Souza SJ, et al.</p>

	Nucleic Acids Research 33(Database Issue): D567-D572. (2005)
達成度評価の方法	テストによる評価
備考	

科目名称	多型解析におけるデータ解釈 (統計学基礎)
区分	必修
講義形態	座学と実習
時間	3時間 (1コマ)
対象者	カリキュラム初級コース受講者 Bio系分野 / IT系分野所属 専門学校または大学学部卒程度で分子生物学・生命科学を学んだ者
対象知識項目	多型解析のための統計学の知識
達成目標	多型解析において解析結果の解釈に必要な統計学の基礎を習得する。
科目概要	<ul style="list-style-type: none"> <li>1 解析結果の解釈 <ul style="list-style-type: none"> <li>1.1 遺伝統計学 <ul style="list-style-type: none"> <li>1.1.1 連鎖解析 <ul style="list-style-type: none"> <li>1.1.1.1 パラメトリック</li> <li>1.1.1.2 ノンパラメトリック</li> <li>1.1.1.3 連鎖不平衡</li> </ul> </li> <li>1.1.2 QTL 解析</li> <li>1.1.3 ハプロタイプ解析</li> </ul> </li> <li>1.2 検定 <ul style="list-style-type: none"> <li>1.2.1 カイ二乗検定</li> <li>1.2.2 ハーディー・ワインバーグ平衡</li> <li>1.2.3 Fisher の正確確率検定</li> </ul> </li> </ul> </li> </ul>
参考文献・資料	ポストゲノム時代の遺伝統計学 鎌谷直之 (編) 羊土社
達成度評価の方法	テストによる評価
備考	

科目名称	プロテオーム解析実験概論
区分	必修
講義形態	座学
時間	3時間 (1コマ)
対象者	カリキュラム初級コース受講者 Bio系分野 / IT系分野所属 専門学校または大学学部卒程度で分子生物学・生命科学を学んだ者
対象知識項目	プロテオーム解析の実験目的、実験原理の知識、プロテオーム解析の実験装置、実験方法の知識
達成目標	プロテオーム解析における実験目的、実験原理、実験装置、実験方法の概要を習得する。
科目概要	<ul style="list-style-type: none"> <li>1 実験の原理・方法 <ul style="list-style-type: none"> <li>1.1 二次元ポリアクリルアミド電気泳動 (2D-PAGE)</li> <li>1.2 プロテアーゼ消化 <ul style="list-style-type: none"> <li>1.2.1 プロテアーゼの基質特異性</li> </ul> </li> <li>1.3 液体クロマトグラフィー (LC)</li> <li>1.4 イオン化法 <ul style="list-style-type: none"> <li>1.4.1 エレクトロスプレーイオン化 (ESI) 法</li> <li>1.4.2 マトリックス支援レーザー脱離イオン化 (MALDI) 法</li> </ul> </li> <li>1.5 質量分析法 <ul style="list-style-type: none"> <li>1.5.1 飛行時間型質量分析計 (TOF - MS)</li> </ul> </li> </ul> </li> </ul>

	1.5.2 四重極型質量分析計 (Q - MS) 1.5.3 イオントラップ型質量分析計 (IT - MS)
参考文献・資料	これからのバイオインフォマティクスのためのバイオ実験入門 実験の原理から理解するバイオの基礎 大藤 道衛 (編集), 高井 貴子 (編集), 高木 利久 羊土社 (2002)  実験医学別冊 決定版プロテオーム解析マニュアル 磯辺俊明, 高橋信弘 (編) 羊土社 (2004)  実験医学別冊 タンパク質実験ハンドブック 分離・精製, 質量分析, 抗体作製, 分子間相互作用解析などの基本原理と最新プロトコール 総集編 竹縄 忠臣 (編) 羊土社 (2003)
達成度評価の方法	テストによる評価
備考	

科目名称	プロテオーム解析におけるデータ解析
区分	必修
講義形態	座学と実習
時間	3 時間 ( 1 コマ )
対象者	カリキュラム初級 / 中級 / 上級コース受講者 Bio 系分野 / IT 系分野所属 専門学校または大学学部卒以上で分子生物学・生命科学を学んだ者
対象知識項目	プロテオーム解析のための検索データベースの知識、プロテオーム解析のための解析ツールの原理の知識、プロテオーム解析のための解析ツールの利用法の知識
達成目標	プロテオーム解析に必要な検索データベースの選択、解析ツールの選択のための知識を習得し、ツールを使いこなす方法を習得する。
科目概要	1 解析方法 ( 解析ツール ) の原理 1.1 PMF ( peptide mass finger printing ) 法 1.2 MS/MS ( tandem MS ) 法 1.3 配列タグ法 1.3.1 アミノ酸配列データベース 1.3.2 仮想のプロテアーゼ消化 1.4 バックグラウンド評価 1.5 スペクトルの帰属 1.6 Mascot  2 解析方法 ( 解析ツール ) の実習 2.1 Mascot
参考文献・資料	実践 バイオインフォマティクス -ゲノム研究のためのコンピュータスキル- Cynthia Gibas (著), Per Jambeck (著), 水島 洋 (翻訳), 明石 浩史 (翻訳), またぬき (翻訳) オライリー・ジャパン (2002)  バイオデータベースとウェブツールの手とり足とり活用法 中村保一, 磯合 敦, 石川 淳 羊土社 (2003)  実験医学別冊 決定版プロテオーム解析マニュアル 磯辺俊明, 高橋信弘 (編) 羊土社 (2004)  実験医学別冊 タンパク質実験ハンドブック 分離・精製, 質量分析, 抗体作製, 分子間相互作用解析などの基本原理と最新プロトコール 総集編

	竹縄 忠臣 (編) 羊土社 (2003)
達成度評価の方法	テストによる評価
備考	

科目名称	プロテオーム解析におけるデータ解釈 (統計学基礎)
区分	必修
講義形態	座学と実習
時間	3 時間 ( 1 コマ )
対象者	カリキュラム初級コース受講者 Bio系分野 / IT系分野所属 専門学校または大学学部卒程度で分子生物学・生命科学を学んだ者
対象知識項目	プロテオーム解析のための統計学の知識
達成目標	プロテオーム解析において解析結果の解釈に必要な統計学の基礎を習得する。
科目概要	1 解析結果の解釈 <ul style="list-style-type: none"> <li>1.1 統計的有意性 <ul style="list-style-type: none"> <li>1.1.1 確率分布 <ul style="list-style-type: none"> <li>1.1.1.1 ポワソン分布、E-value、p-value</li> </ul> </li> <li>1.1.2 標準偏差 <ul style="list-style-type: none"> <li>1.1.2.1 Z score</li> </ul> </li> <li>1.1.3 尤度</li> </ul> </li> <li>1.2 質量誤差の評価</li> <li>1.3 他の解析法との比較</li> </ul>
参考文献・資料	実験医学別冊 決定版プロテオーム解析マニュアル 磯辺俊明, 高橋信弘 (編) 羊土社 (2004)  実験医学別冊 タンパク質実験ハンドブック 分離・精製, 質量分析, 抗体作製, 分子間相互作用解析などの基本原理と最新プロトコール 総集編 竹縄 忠臣 (編) 羊土社 (2003)
達成度評価の方法	テストによる評価
備考	

科目名称	タンパク質間ネットワーク解析実験概論
区分	必修
講義形態	座学
時間	3 時間 ( 1 コマ )
対象者	カリキュラム初級コース受講者 Bio系分野 / IT系分野所属 専門学校または大学学部卒程度で分子生物学・生命科学を学んだ者
対象知識項目	タンパク質間ネットワーク解析の実験目的、実験原理の知識、タンパク質間ネットワーク解析の実験装置、実験方法の知識
達成目標	タンパク質間ネットワーク解析における実験目的、実験原理、実験装置、実験方法の概要を習得する。
科目概要	1 実験の原理・方法 <ul style="list-style-type: none"> <li>1.1 タンパク質間相互作用</li> <li>1.2 代謝パスウェイ</li> <li>1.3 酵母 two-hybrid 法 <ul style="list-style-type: none"> <li>1.3.1 cDNA ライブラリーからのスクリーニング</li> <li>1.3.2 転写調節因子</li> <li>1.3.3 レポーター遺伝子</li> </ul> </li> <li>1.4 ブルダウン法 <ul style="list-style-type: none"> <li>1.4.1 質量分析法</li> </ul> </li> </ul>

	1.4.2 ウェスタンブロット法 1.5 プロテインチップ 1.5.1 質量分析法
参考文献・資料	これからのバイオインフォマティクスのためのバイオ実験入門 実験の原理から理解するバイオの基礎 大藤 道衛 (編集), 高井 貴子 (編集), 高木 利久 羊土社 (2002)  実験医学別冊 決定版プロテオーム解析マニュアル 磯辺俊明, 高橋信弘 (編) 羊土社 (2004)  実験医学別冊 タンパク質実験ハンドブック 分離・精製, 質量分析, 抗体作製, 分子間相互作用解析などの基本原理と最新プロトコール 総集編 竹縄 忠臣 (編) 羊土社 (2003)
達成度評価の方法	テストによる評価
備考	

科目名称	タンパク質間ネットワーク解析におけるデータ解析
区分	必修
講義形態	座学と実習
時間	3 時間 ( 1 コマ )
対象者	カリキュラム初級 / 中級 / 上級コース受講者 Bio 系分野 / IT 系分野所属 専門学校または大学学部卒以上で分子生物学・生命科学を学んだ者
対象知識項目	タンパク質間ネットワーク解析のための検索データベースの知識、タンパク質間ネットワーク解析のための解析ツールの原理の知識、タンパク質間ネットワーク解析のための解析ツールの利用法の知識
達成目標	タンパク質間ネットワーク解析に必要な検索データベースの選択、解析ツールの選択のための知識を習得し、ツールを使いこなす方法を習得する。
科目概要	1 解析方法 ( 解析ツール ) の原理 1.1 PMF ( peptide mass finger printing ) 法 1.2 MS/MS ( tandem MS ) 法 1.3 配列タグ法 1.3.1 アミノ酸配列データベース 1.3.2 仮想のプロテアーゼ消化 1.4 バックグラウンド評価 1.5 スペクトルの帰属 1.5.1 決定木 1.6 自己組織化マップ ( SOM ) 1.7 複合グラフモデル 1.8 パスウェイ記述法 ( XML、ペトリネット ) 1.9 Mascot  2 解析方法 ( 解析ツール ) の実習 2.1 Mascot
参考文献・資料	ゲノムネットのデータベース利用法 第 3 版 金久 実 ( 編 ) 共立出版 ( 2002 )  実践 バイオインフォマティクス -ゲノム研究のためのコンピュータスキル- Cynthia Gibas ( 著 ), Per Jambeck ( 著 ), 水島 洋 ( 翻訳 ), 明石 浩史 ( 翻訳 ), ま たぬき ( 翻訳 ) オライリー・ジャパン ( 2002 )

	<p>バイオデータベースとウェブツールの手とり足とり活用法 中村保一，磯合 敦，石川 淳 羊土社（2003）</p> <p>実験医学別冊 決定版プロテオーム解析マニュアル 磯辺俊明，高橋信弘（編）羊土社（2004）</p> <p>実験医学別冊 タンパク質実験ハンドブック 分離・精製，質量分析，抗体作製，分子間相互作用解析などの基本原理と最新プロトコール 総集編 竹縄 忠臣（編）羊土社（2003）</p>
達成度評価の方法	テストによる評価
備考	

科目名称	タンパク質間ネットワーク解析におけるデータ解釈（統計学基礎）
区分	必修
講義形態	座学と実習
時間	3時間（1コマ）
対象者	カリキュラム初級コース受講者 Bio系分野 / IT系分野所属 専門学校または大学学部卒程度で分子生物学・生命科学を学んだ者
対象知識項目	タンパク質間ネットワーク解析のための統計学の知識
達成目標	タンパク質間ネットワーク解析において解析結果の解釈に必要な統計学の基礎を習得する。
科目概要	<p>1 解析結果の解釈</p> <p>1.1 統計的有意性</p> <p>1.1.1 確率分布</p> <p>1.1.1.1 ポワソン分布、E-value、p-value</p> <p>1.1.2 標準偏差</p> <p>1.1.2.1 Z score</p> <p>1.1.3 尤度</p> <p>1.2 質量誤差の評価</p> <p>1.3 他の解析法との比較</p>
参考文献・資料	<p>実験医学別冊 決定版プロテオーム解析マニュアル 磯辺俊明，高橋信弘（編）羊土社（2004）</p> <p>実験医学別冊 タンパク質実験ハンドブック 分離・精製，質量分析，抗体作製，分子間相互作用解析などの基本原理と最新プロトコール 総集編 竹縄 忠臣（編）羊土社（2003）</p>
達成度評価の方法	テストによる評価
備考	

科目名称	アノテーター序説
区分	必修
講義形態	座学
時間	3時間（1コマ）
対象者	カリキュラム初級 / 中級 / 上級コース受講者 Bio系分野 / IT系分野所属 専門学校または大学学部卒以上で分子生物学・生命科学を学んだ者
対象知識項目	データベースの国際的な基準への準拠、データの信頼度を高める一貫性、継続性、確実性、生命科学の急速な進展と並走する前進性、実務の倫理性
達成目標	データベースの国際的な基準への準拠、データの信頼度を高める一貫性・継続性・確実性、生命科学の急速な進展と並走する前進性、実務の倫理性について習

	得する。
<b>科目概要</b>	<p>1 データベースの国際的な基準への準拠</p> <p>1.1 データベースの最新情報</p> <p>1.2 国内外の研究機関との連携</p> <p>1.3 最新文献の理解</p> <p>2 データの信頼度を高める一貫性、継続性、確実性</p> <p>2.1 統一ルールの遵守</p> <p>2.2 統一ルールの作成</p> <p>3 生命科学の急速な進展と並走する前進性</p> <p>3.1 最新文献の理解</p> <p>4 実務の倫理性</p> <p>4.1 個人情報の保護</p> <p>4.2 遺伝子取り扱い倫理規定解析方法（解析ツール）の原理</p>
<b>参考文献・資料</b>	未定
<b>達成度評価の方法</b>	テストによる評価
<b>備考</b>	

<b>科目名称</b>	分子生物学概論
<b>区分</b>	必修
<b>講義形態</b>	座学
<b>時間</b>	3時間（1コマ）
<b>対象者</b>	カリキュラム中級／上級コース受講者 Bio系分野／IT系分野所属 大学修士課程以上で分子生物学・生命科学を学んだ者
<b>対象知識項目</b>	分子生物学の知識
<b>達成目標</b>	データ解釈のために必要な分子生物学の概要、文献読解、Gene ontology の利用について習得する。
<b>科目概要</b>	<p>1 分子生物学用語の知識</p> <p>2 文献読解</p> <p>3 Gene ontology</p>
<b>参考文献・資料</b>	<p>ゲノム2 新しい生命情報システムへのアプローチ T.A. Brown (著), 村松 正実 (翻訳) メディカルサイエンスインターナショナル (2003)</p> <p>細胞の分子生物学 第4版 Bruce Alberts (著), 中村 桂子 (翻訳), 松原 謙一 (翻訳) ニュートンプレス (2004)</p>
<b>達成度評価の方法</b>	テストによる評価
<b>備考</b>	

<b>科目名称</b>	配列解析実験特論
<b>区分</b>	必修
<b>講義形態</b>	座学
<b>時間</b>	3時間（1コマ）
<b>対象者</b>	カリキュラム中級／上級コース受講者 Bio系分野／IT系分野所属 大学修士課程以上で分子生物学・生命科学を学んだ者

<b>対象知識項目</b>	配列解析の実験目的、実験原理の知識、配列解析の実験装置、実験方法の知識
<b>達成目標</b>	配列解析における実験目的、実験原理、実験装置、実験方法を習得する。特に生データの精度まで判断できる方法を習得する。
<b>科目概要</b>	<p>1 基礎的な背景知識（基礎のため省略可）</p> <p>1.1 クローニングの方法</p> <p>1.2 シーケンスの原理と方法</p> <p>2 実験の原理・方法</p> <p>2.1 Base call の原理と方法</p> <p>2.2 Assembly の原理と方法</p> <p>2.3 ショットガンシーケンス</p>
<b>参考文献・資料</b>	これからのバイオインフォマティクスのためのバイオ実験入門 実験の原理から理解するバイオの基礎 大藤 道衛（編集）、高井 貴子（編集）、高木 利久 羊土社（2002）
<b>達成度評価の方法</b>	テストによる評価
<b>備考</b>	

<b>科目名称</b>	配列解析におけるデータ解釈(統計学)
<b>区分</b>	必修
<b>講義形態</b>	座学と実習
<b>時間</b>	3時間（1コマ）
<b>対象者</b>	カリキュラム中級/上級コース受講者 Bio系分野/IT系分野所属 大学修士課程以上で分子生物学・生命科学を学んだ者
<b>対象知識項目</b>	配列解析のための統計学の知識
<b>達成目標</b>	配列解析において解析結果の解釈に必要な統計学を習得する。
<b>科目概要</b>	<p>1 解析結果の解釈</p> <p>1.1 統計的有意性</p> <p>1.1.1 確率分布</p> <p>1.1.1.1 ポワソン分布、E-value、p-value</p> <p>1.1.2 標準偏差</p> <p>1.1.2.1 Z score</p> <p>1.1.3 尤度</p>
<b>参考文献・資料</b>	<p>バイオインフォマティクス ゲノム配列から機能解析へ 岡崎 康司（翻訳）、坊農 秀雅（翻訳）、David W. Mount メディカル・サイエンス・インターナショナル（2002）</p> <p>Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins (Methods of Biochemical Analysis, V. 43) Edited by: Andreas D. Baxevanis &amp; B. F. Francis Ouellette WILEY-INTERSCIENCE (2001)</p> <p>BLAST: at the core of a powerful and diverse set of sequence analysis tools. Scott McGinnis and Thomas L. Madden Nucleic Acids Research 2004, 32(Web Server issue):W20-W25</p>
<b>達成度評価の方法</b>	テストによる評価
<b>備考</b>	

<b>科目名称</b>	配列解析におけるデータ解釈（総合判断）
<b>区分</b>	必修

講義形態	座学と実習
時間	3 時間 ( 1 コマ )
対象者	カリキュラム中級コース受講者 Bio 系分野 / IT 系分野所属 大学修士課程以上で分子生物学・生命科学を学んだ者
対象知識項目	配列解析のための総合判断・産業応用の知識
達成目標	配列解析において解析結果の解釈に必要な機能推定、複数の解析方法による結果の比較・検討、機能の応用利用について習得する。
科目概要	1 解析結果の解釈 <ul style="list-style-type: none"> <li>1.1 遺伝子機能のアノテーション <ul style="list-style-type: none"> <li>1.1.1 検索結果の解釈 <ul style="list-style-type: none"> <li>1.1.1.1 既存データベースのアノテーションの読み方 <ul style="list-style-type: none"> <li>1.1.1.1.1 実験事実か予測によるものか</li> </ul> </li> </ul> </li> <li>1.1.2 Gene ontology <ul style="list-style-type: none"> <li>1.1.2.1 Gene Ontology Annotation</li> <li>1.1.2.2 go-dev</li> </ul> </li> </ul> </li> <li>1.2 総合的解釈 <ul style="list-style-type: none"> <li>1.2.1 機能推定</li> <li>1.2.2 複数の検索・予測結果の比較・検討</li> </ul> </li> </ul>
参考文献・資料	生命情報学キーノート D.R. Westhead, J.H. Parish, R.M. Twyman 監訳：池尾一穂，鈴木善幸，五條堀 孝 シュプリンガーフェアラーク東京 ( 2003 )  バイオインフォマティクス 応用生命科学シリーズ 9 美宅成樹，榊 佳之 東京化学同人 ( 2003 )  バイオインフォマティクス ゲノム配列から機能解析へ 岡崎 康司 ( 翻訳 )，坊農 秀雅 ( 翻訳 )，David W. Mount メディカル・サイエンス・インターナショナル ( 2002 )
達成度評価の方法	テストによる評価
備考	

科目名称	発現プロファイル解析実験特論
区分	必修
講義形態	座学
時間	3 時間 ( 1 コマ )
対象者	カリキュラム中級 / 上級コース受講者 Bio 系分野 / IT 系分野所属 大学修士課程以上で分子生物学・生命科学を学んだ者
対象知識項目	発現プロファイル解析の実験目的、実験原理の知識、発現プロファイル解析の実験装置、実験方法の知識
達成目標	発現プロファイル解析における実験目的、実験原理、実験装置、実験方法を習得する。特に生データの精度まで判断できる方法を習得する。
科目概要	1 基礎的な背景知識 ( 基礎のため省略可 ) <ul style="list-style-type: none"> <li>1.1 遺伝子発現のしくみ</li> <li>1.2 ハイブリダイゼーションの原理</li> <li>1.3 蛍光ラベリングの原理</li> </ul> 2 実験の原理・方法 <ul style="list-style-type: none"> <li>2.1 DNAマイクロアレイ、DNAチップの種類 <ul style="list-style-type: none"> <li>2.1.1 チップ型アレイ ( GeneChip )</li> <li>2.1.2 スポット型アレイ ( cDNA 型アレイ )</li> </ul> </li> </ul>

	<p>3 実験準備の知識</p> <p>3.1 調整試料</p> <p>3.1.1 cDNA ライブラリー</p> <p>3.1.2 RNA 抽出</p> <p>3.1.3 PCR</p> <p>3.2 試薬</p> <p>3.2.1 蛍光色素 (Cy3, Cy5)</p> <p>3.3 機材</p> <p>3.3.1 スキャナー</p> <p>3.3.2 アレイ / チップ上の遺伝子の既知情報</p>
参考文献・資料	<p>これからのバイオインフォマティクスのためのバイオ実験入門 実験の原理から理解するバイオの基礎</p> <p>大藤 道衛 (編集), 高井 貴子 (編集), 高木 利久 羊土社 (2002)</p> <p>DNA マイクロアレイ実践マニュアル</p> <p>林崎良英(監修), 岡崎康司(編) 羊土社 (2000)</p>
達成度評価の方法	テストによる評価
備考	

科目名称	発現プロファイル解析におけるデータ解釈 (統計学)
区分	必修
講義形態	座学と実習
時間	3 時間 (1 コマ)
対象者	カリキュラム中級 / 上級コース受講者 Bio 系分野 / IT 系分野所属 大学修士課程以上で分子生物学・生命科学を学んだ者
対象知識項目	発現プロファイル解析のための統計学の知識
達成目標	発現プロファイル解析において解析結果の解釈に必要な統計学を習得する。
科目概要	<p>1 解析結果の解釈</p> <p>1.1 クラスタ解析</p> <p>1.1.1 クラスタ化法の種類</p> <p>1.1.1.1 階層型クラスタ化法、K-means クラスタ化法、自己組織化マップ (SOM)</p> <p>1.1.2 相関解析</p> <p>1.1.3 多変量解析</p> <p>2.2 多変量解析ソフト</p> <p>2.2.1 R</p> <p>2.2.2 GeneSpring</p>
参考文献・資料	R による統計解析の基礎 Computer in Education and Research 7 中澤 港 ピアソンエデュケーション (2003)
達成度評価の方法	テストによる評価
備考	

科目名称	発現プロファイル解析におけるデータ解釈 (総合判断)
区分	必修
講義形態	座学と実習
時間	3 時間 (1 コマ)
対象者	カリキュラム中級コース受講者 Bio 系分野 / IT 系分野所属

	大学修士課程以上で分子生物学・生命科学を学んだ者
対象知識項目	発現プロファイル解析のための総合判断・産業応用の知識
達成目標	発現プロファイル解析において解析結果の解釈に必要な機能推定、複数の解析方法による結果の比較・検討、機能の応用利用について習得する。
科目概要	<ol style="list-style-type: none"> <li>1. 解析結果の解釈 <ol style="list-style-type: none"> <li>1.1 機能推定 <ol style="list-style-type: none"> <li>1.1.1 組織特異的発現</li> </ol> </li> <li>1.2 遺伝子ネットワーク解析 <ol style="list-style-type: none"> <li>1.2.1 プーリアンモデル</li> <li>1.2.2 ベイジアンモデル</li> <li>1.2.3 S-system</li> </ol> </li> <li>1.3 システムバイオロジー</li> <li>1.4 解析結果の比較・検討</li> <li>1.5 疾患等との関連</li> </ol> </li> </ol>
参考文献・資料	<p>バイオインフォマティクス ゲノム配列から機能解析へ 岡崎 康司 (翻訳), 坊農 秀雅 (翻訳), David W. Mount メディカル・サイエンス・インターナショナル (2002)</p> <p>統合ゲノミクスのためのマイクロアレイデータアナリシス Kohane, I.S., Kho, A.T., Butte, A.J. (著), 星田有人 (訳) シュプリンガーフェアラーク東京 (2004)</p> <p>DNA マイクロアレイデータ解析入門 Knudsen, S. (著), 塩島聡, 松本治, 辻本豪三 (監訳) 羊土社 (2002)</p>
達成度評価の方法	テストによる評価
備考	

科目名称	多型解析実験特論
区分	必修
講義形態	座学
時間	3 時間 ( 1 コマ )
対象者	カリキュラム中級 / 上級コース受講者 Bio 系分野 / IT 系分野所属 大学修士課程以上で分子生物学・生命科学を学んだ者
対象知識項目	多型解析の実験目的、実験原理の知識、多型解析の実験装置、実験方法の知識
達成目標	多型解析における実験目的、実験原理、実験装置、実験方法を習得する。特に生データの精度まで判断できる方法を習得する。
科目概要	<ol style="list-style-type: none"> <li>1 基礎的な背景知識 <ol style="list-style-type: none"> <li>1.1 遺伝子多型の種類 <ol style="list-style-type: none"> <li>1.1.1 RFLP</li> <li>1.1.2 VNTR</li> <li>1.1.3 マイクロサテライト</li> <li>1.1.4 SNP</li> </ol> </li> <li>1.2 SNP の種類</li> <li>1.3 SNP と疾患 <ol style="list-style-type: none"> <li>1.3.1 単一遺伝子疾患</li> <li>1.3.2 多因子疾患</li> </ol> </li> <li>1.4 SNP と薬物応答性</li> </ol> </li> <li>2 実験の原理・方法 <ol style="list-style-type: none"> <li>2.1 SNP の同定法 <ol style="list-style-type: none"> <li>2.1.1 シーケンス</li> <li>2.1.2 Assembly</li> </ol> </li> </ol> </li> </ol>

	2.1.3 Invader 法 2.1.4 ICAM 法 2.1.5 DNA マイクロアレイ、DNA チップ
参考文献・資料	これからのバイオインフォマティクスのためのバイオ実験入門 実験の原理から理解するバイオの基礎 大藤 道衛 (編集), 高井 貴子 (編集), 高木 利久 羊土社 (2002)
達成度評価の方法	テストによる評価
備考	

科目名称	多型解析におけるデータ解釈 (統計学)
区分	必修
講義形態	座学と実習
時間	3 時間 (1 コマ)
対象者	カリキュラム中級 / 上級コース受講者 Bio 系分野 / IT 系分野所属 大学修士課程以上で分子生物学・生命科学を学んだ者
対象知識項目	多型解析のための統計学の知識
達成目標	多型解析において解析結果の解釈に必要な統計学を習得する。
科目概要	1 解析結果の解釈 1.1 遺伝統計学 1.1.1 連鎖解析 1.1.1.1 パラメトリック 1.1.1.2 ノンパラメトリック 1.1.1.3 連鎖不平衡 1.1.2 QTL 解析 1.1.3 ハプロタイプ解析 1.2 検定 1.2.1 カイ二乗検定 1.2.2 ハーディー・ワインバーグ平衡 1.2.3 Fisher の正確確率検定
参考文献・資料	ポストゲノム時代の遺伝統計学 鎌谷直之 (編) 羊土社
達成度評価の方法	テストによる評価
備考	

科目名称	多型解析におけるデータ解釈 (総合判断)
区分	必修
講義形態	座学と実習
時間	3 時間 (1 コマ)
対象者	カリキュラム中級コース受講者 Bio 系分野 / IT 系分野所属 大学修士課程以上で分子生物学・生命科学を学んだ者
対象知識項目	多型解析のための総合判断・産業応用の知識
達成目標	多型解析において解析結果の解釈に必要な機能推定、複数の解析方法による結果の比較・検討、機能の応用利用について習得する。
科目概要	1 解析結果の解釈 1.1 疾患との関連 1.1.1 遺伝子多型を利用した疾患関連遺伝子探索 1.2 薬剤応答性との関連 1.3 テーラーメイド医療

<b>参考文献・資料</b>	<p>ヒト遺伝子アノテーションの統合データベース ゲノム医学 Vol.4, No.4 山崎千里、今西規 メディカルレビュー社 (2004)</p> <p>遺伝子情報統合データベース H-Invitational Database 藤井康之、今西規、五條堀孝 バイオ高性能機器新技術利用マニュアル、蛋白質核酸酵素増刊号 小原収、谷口寿章、市川哲生、猪飼篤、編 (2004)</p> <p>Integrative Annotation of 21,037 Human Genes Validated by Full-Length cDNA Clones. Imanishi T, Itoh T, Suzuki Y, O'Donovan C, Fukuchi S, Koyanagi KO, Barrero RA, Tamura T, Yamaguchi-Kabata Y, Tanino M, et al. PLoS Biology 2: 856-875. (2004)</p> <p>The Human Anatomic Gene Expression Library (H-ANGEL), the H-Inv Integrative Display of Human Gene Expression across Disparate Technologies and Platforms. Tanino M, Debily MA, Tamura T, Hishiki T, Ogasawara O, Murakawa K, Kawamoto S, Itoh K, Watanabe S, de Souza SJ, et al. Nucleic Acids Research 33(Database Issue): D567-D572. (2005)</p>
<b>達成度評価の方法</b>	テストによる評価
<b>備考</b>	

<b>科目名称</b>	プロテオーム解析実験特論
<b>区分</b>	必修
<b>講義形態</b>	座学
<b>時間</b>	3時間 (1コマ)
<b>対象者</b>	カリキュラム中級/上級コース受講者 Bio系分野/IT系分野所属 大学修士課程以上で分子生物学・生命科学を学んだ者
<b>対象知識項目</b>	プロテオーム解析の実験目的、実験原理の知識、プロテオーム解析の実験装置、実験方法の知識
<b>達成目標</b>	プロテオーム解析における実験目的、実験原理、実験装置、実験方法を習得する。特に生データの精度まで判断できる方法を習得する。
<b>科目概要</b>	<p>1 実験の原理・方法</p> <p>1.1 二次元ポリアクリルアミド電気泳動 (2D-PAGE)</p> <p>1.2 プロテアーゼ消化</p> <p>1.2.1 プロテアーゼの基質特異性</p> <p>1.3 液体クロマトグラフィー (LC)</p> <p>1.4 イオン化法</p> <p>1.4.1 エレクトロスプレーイオン化 (ESI) 法</p> <p>1.4.2 マトリックス支援レーザー脱離イオン化 (MALDI) 法</p> <p>1.5 質量分析法</p> <p>1.5.1 飛行時間型質量分析計 (TOF-MS)</p> <p>1.5.2 四重極型質量分析計 (Q-MS)</p> <p>1.5.3 イオントラップ型質量分析計 (IT-MS)</p>
<b>参考文献・資料</b>	<p>これからのバイオインフォマティクスのためのバイオ実験入門 実験の原理から理解するバイオの基礎 大藤 道衛 (編集), 高井 貴子 (編集), 高木 利久 羊土社 (2002)</p> <p>実験医学別冊 決定版プロテオーム解析マニュアル 磯辺俊明, 高橋信弘 (編) 羊土社 (2004)</p>

	実験医学別冊 タンパク質実験ハンドブック 分離・精製，質量分析，抗体作製，分子間相互作用解析などの基本原理と最新プロトコール 総集編 竹縄 忠臣（編） 羊土社（2003）
達成度評価の方法	テストによる評価
備考	

科目名称	プロテオーム解析におけるデータ解釈（統計学）
区分	必修
講義形態	座学と実習
時間	3時間（1コマ）
対象者	カリキュラム中級/上級コース受講者 Bio系分野/IT系分野所属 大学修士課程以上で分子生物学・生命科学を学んだ者
対象知識項目	プロテオーム解析のための統計学の知識
達成目標	プロテオーム解析において解析結果の解釈に必要な統計学を習得する。
科目概要	1 解析結果の解釈 1.1 統計的有意性 1.1.1 確率分布 1.1.1.1 ポワソン分布、E-value、p-value 1.1.2 標準偏差 1.1.2.1 Z score 1.1.3 尤度 1.2 質量誤差の評価 1.3 他の解析法との比較
参考文献・資料	実験医学別冊 決定版プロテオーム解析マニュアル 磯辺俊明，高橋信弘（編）羊土社（2004）  実験医学別冊 タンパク質実験ハンドブック 分離・精製，質量分析，抗体作製，分子間相互作用解析などの基本原理と最新プロトコール 総集編 竹縄 忠臣（編） 羊土社（2003）
達成度評価の方法	テストによる評価
備考	

科目名称	プロテオーム解析におけるデータ解釈（総合判断）
区分	必修
講義形態	座学と実習
時間	3時間（1コマ）
対象者	カリキュラム中級コース受講者 Bio系分野/IT系分野所属 大学修士課程以上で分子生物学・生命科学を学んだ者
対象知識項目	プロテオーム解析のための総合判断・産業応用の知識
達成目標	プロテオーム解析において解析結果の解釈に必要な機能推定、複数の解析方法による結果の比較・検討、機能の応用利用について習得する。
科目概要	1 解析結果の解釈 1.1 質量誤差の評価 1.2 翻訳後修飾の同定 1.3 タンパク質間相互作用 1.4 遺伝子ネットワーク解析 1.4.1 プーリアンモデル 1.4.2 ベイジアンモデル

	1.4.3 S-system 1.5 システムバイオロジー 1.6 総合的解釈 1.6.1 機能推定 1.6.2 他の解析法との比較 1.6.3 疾患等との関連 1.6.3.1 疾患原因遺伝子 1.6.3.2 診断マーカー
参考文献・資料	実験医学別冊 決定版プロテオーム解析マニュアル 磯辺俊明, 高橋信弘 (編) 羊土社 (2004)  実験医学別冊 タンパク質実験ハンドブック 分離・精製, 質量分析, 抗体作製, 分子間相互作用解析などの基本原理と最新プロトコール 総集編 竹縄 忠臣 (編) 羊土社 (2003)
達成度評価の方法	テストによる評価
備考	

科目名称	タンパク質間ネットワーク解析実験特論
区分	必修
講義形態	座学
時間	3時間 (1コマ)
対象者	カリキュラム中級/上級コース受講者 Bio系分野/IT系分野所属 大学修士課程以上で分子生物学・生命科学を学んだ者
対象知識項目	タンパク質間ネットワーク解析の実験目的、実験原理の知識、タンパク質間ネットワーク解析の実験装置、実験方法の知識
達成目標	タンパク質間ネットワーク解析における実験目的、実験原理、実験装置、実験方法を習得する。特に生データの精度まで判断できる方法を習得する。
科目概要	1 実験の原理・方法 1.1 タンパク質間相互作用 1.2 代謝パスウェイ 1.3 酵母 two-hybrid 法 1.3.1 cDNA ライブラリーからのスクリーニング 1.3.2 転写調節因子 1.3.3 レポーター遺伝子 1.4 プルダウン法 1.4.1 質量分析法 1.4.2 ウェスタンブロット法 1.5 プロテインチップ 1.5.1 質量分析法
参考文献・資料	これからのバイオインフォマティクスのためのバイオ実験入門 実験の原理から理解するバイオの基礎 大藤 道衛 (編集), 高井 貴子 (編集), 高木 利久 羊土社 (2002)  実験医学別冊 決定版プロテオーム解析マニュアル 磯辺俊明, 高橋信弘 (編) 羊土社 (2004)  実験医学別冊 タンパク質実験ハンドブック 分離・精製, 質量分析, 抗体作製, 分子間相互作用解析などの基本原理と最新プロトコール 総集編 竹縄 忠臣 (編) 羊土社 (2003)
達成度評価の方法	テストによる評価

備考	
科目名称	タンパク質間ネットワーク解析におけるデータ解釈(統計学)
区分	必修
講義形態	座学と実習
時間	3時間(1コマ)
対象者	カリキュラム中級/上級コース受講者 Bio系分野/IT系分野所属 大学修士課程以上で分子生物学・生命科学を学んだ者
対象知識項目	タンパク質間ネットワーク解析のための統計学の知識
達成目標	タンパク質間ネットワーク解析において解析結果の解釈に必要な統計学を習得する。
科目概要	1 解析結果の解釈 <ul style="list-style-type: none"> <li>1.1 統計的有意性 <ul style="list-style-type: none"> <li>1.1.1 確率分布 <ul style="list-style-type: none"> <li>1.1.1.1 ポワソン分布、E-value、p-value</li> </ul> </li> <li>1.1.2 標準偏差 <ul style="list-style-type: none"> <li>1.1.2.1 Z score</li> </ul> </li> <li>1.1.3 尤度</li> </ul> </li> <li>1.2 質量誤差の評価</li> <li>1.3 他の解析法との比較</li> </ul>
参考文献・資料	実験医学別冊 決定版プロテオーム解析マニュアル 磯辺俊明, 高橋信弘(編) 羊土社(2004)  実験医学別冊 タンパク質実験ハンドブック 分離・精製, 質量分析, 抗体作製, 分子間相互作用解析などの基本原理と最新プロトコール 総集編 竹縄 忠臣(編) 羊土社(2003)
達成度評価の方法	テストによる評価
備考	

科目名称	タンパク質間ネットワーク解析におけるデータ解釈(総合判断)
区分	必修
講義形態	座学と実習
時間	3時間(1コマ)
対象者	カリキュラム中級コース受講者 Bio系分野/IT系分野所属 大学修士課程以上で分子生物学・生命科学を学んだ者
対象知識項目	タンパク質間ネットワーク解析のための総合判断・産業応用の知識
達成目標	タンパク質間ネットワーク解析において解析結果の解釈に必要な機能推定、複数の解析方法による結果の比較・検討、機能の応用利用について習得する。
科目概要	1 解析結果の解釈 <ul style="list-style-type: none"> <li>1.1 質量誤差の評価</li> <li>1.2 翻訳後修飾の同定</li> <li>1.3 タンパク質間相互作用</li> <li>1.4 遺伝子ネットワーク解析 <ul style="list-style-type: none"> <li>1.4.1 プーリアンモデル</li> <li>1.4.2 ベイジアンモデル</li> <li>1.4.3 S-system</li> </ul> </li> <li>1.5 システムバイオロジー</li> <li>1.6 総合的解釈 <ul style="list-style-type: none"> <li>1.6.1 機能推定</li> <li>1.6.2 他の解析法との比較</li> </ul> </li> </ul>

	<p>1.6.3 疾患等との関連</p> <p>1.6.3.1 疾患原因遺伝子</p> <p>1.6.3.2 診断マーカー</p>
参考文献・資料	<p>実験医学別冊 決定版プロテオーム解析マニュアル 磯辺俊明, 高橋信弘 (編) 羊土社 (2004)</p> <p>実験医学別冊 タンパク質実験ハンドブック 分離・精製, 質量分析, 抗体作製, 分子間相互作用解析などの基本原理と最新プロトコール 総集編 竹縄 忠臣 (編) 羊土社 (2003)</p>
達成度評価の方法	テストによる評価
備考	

科目名称	配列解析におけるデータ解釈 (総合判断・産業応用)
区分	必修
講義形態	座学と実習
時間	3 時間 (1 コマ)
対象者	カリキュラム上級コース受講者 Bio系分野 / IT系分野所属 大学博士課程以上で分子生物学・生命科学を学んだ者
対象知識項目	配列解析のための総合判断・産業応用の知識
達成目標	配列解析において解析結果の解釈に必要な機能推定、複数の解析方法による結果の比較・検討、機能の応用利用について習得する。
科目概要	<p>1 解析結果の解釈</p> <p>1.1 遺伝子機能のアノテーション</p> <p>1.1.1 検索結果の解釈</p> <p>1.1.1.1 既存データベースのアノテーションの読み方</p> <p>1.1.1.1.1 実験事実か予測によるものか</p> <p>1.1.2 Gene ontology</p> <p>1.1.2.1 Gene Ontology Annotation</p> <p>1.1.2.2 go-dev</p> <p>1.2 総合的解釈</p> <p>1.2.1 機能推定</p> <p>1.2.2 複数の検索・予測結果の比較・検討</p> <p>1.2.3 疾患等との関連 (実験サンプルのプロファイルの熟知から解析結果の解釈、疾患等遺伝子機能の応用)</p>
参考文献・資料	<p>生命情報学キーノート D.R. Westhead, J.H. Parish, R.M. Twyman 監訳: 池尾一穂, 鈴木善幸, 五條堀 孝 シュプリンガーフェアラーク東京 (2003)</p> <p>バイオインフォマティクス 応用生命科学シリーズ9 美宅成樹, 榊 佳之 東京化学同人 (2003)</p> <p>バイオインフォマティクス ゲノム配列から機能解析へ 岡崎 康司 (翻訳), 坊農 秀雅 (翻訳), David W. Mount メディカル・サイエンス・インターナショナル (2002)</p>
達成度評価の方法	テストによる評価
備考	

科目名称	発現プロファイル解析におけるデータ解釈（総合判断・産業応用）
区分	必修
講義形態	座学と実習
時間	3時間（1コマ）
対象者	カリキュラム上級コース受講者 Bio系分野/IT系分野所属 大学博士課程以上で分子生物学・生命科学を学んだ者
対象知識項目	発現プロファイル解析のための総合判断・産業応用の知識
達成目標	発現プロファイル解析において解析結果の解釈に必要な機能推定、複数の解析方法による結果の比較・検討、機能の応用利用について習得する。
科目概要	<ol style="list-style-type: none"> <li>1. 解析結果の解釈 <ol style="list-style-type: none"> <li>1.1 機能推定 <ol style="list-style-type: none"> <li>1.1.1 組織特異的発現</li> </ol> </li> <li>1.2 遺伝子ネットワーク解析 <ol style="list-style-type: none"> <li>1.2.1 ブーリアンモデル</li> <li>1.2.2 ベイジアンモデル</li> <li>1.2.3 S-system</li> </ol> </li> <li>1.3 システムバイオロジー</li> <li>1.4 解析結果の比較・検討</li> <li>1.5 疾患等との関連</li> <li>1.6 医療及び産業応用</li> </ol> </li> </ol>
参考文献・資料	<p>バイオインフォマティクス ゲノム配列から機能解析へ 岡崎 康司（翻訳）、坊農 秀雅（翻訳）、David W. Mount メディカル・サイエンス・インターナショナル（2002）</p> <p>統合ゲノミクスのためのマイクロアレイデータアナリシス Kohane, I.S., Kho, A.T., Butte, A.J. (著), 星田有人(訳) シュプリンガーフェアラーク東京（2004）</p> <p>DNA マイクロアレイデータ解析入門 Knudsen, S. (著), 塩島聡, 松本治, 辻本豪三（監訳） 羊土社（2002）</p>
達成度評価の方法	テストによる評価
備考	

科目名称	多型解析におけるデータ解釈（総合判断・産業応用）
区分	必修
講義形態	座学と実習
時間	3時間（1コマ）
対象者	カリキュラム上級コース受講者 Bio系分野/IT系分野所属 大学博士課程以上で分子生物学・生命科学を学んだ者
対象知識項目	多型解析のための総合判断・産業応用の知識
達成目標	多型解析において解析結果の解釈に必要な機能推定、複数の解析方法による結果の比較・検討、機能の応用利用について習得する。
科目概要	<ol style="list-style-type: none"> <li>1 解析結果の解釈 <ol style="list-style-type: none"> <li>1.1 疾患との関連 <ol style="list-style-type: none"> <li>1.1.1 遺伝子多型を利用した疾患関連遺伝子探索</li> </ol> </li> <li>1.2 薬剤応答性との関連</li> <li>1.3 テーラーメイド医療</li> <li>1.4 産業応用</li> </ol> </li> </ol>
参考文献・資料	ヒト遺伝子アノテーションの統合データベース ゲノム医学 Vol.4, No.4

	<p>山崎千里、今西規 メディカルレビュー社 (2004)</p> <p>遺伝子情報統合データベース H-Invitational Database 藤井康之、今西規、五條堀孝 バイオ高性能機器新技術利用マニュアル、蛋白質核酸酵素増刊号 小原収、谷口寿章、市川哲生、猪飼篤、編 (2004)</p> <p>Integrative Annotation of 21,037 Human Genes Validated by Full-Length cDNA Clones. Imanishi T, Itoh T, Suzuki Y, O'Donovan C, Fukuchi S, Koyanagi KO, Barrero RA, Tamura T, Yamaguchi-Kabata Y, Tanino M, et al. PLoS Biology 2: 856-875. (2004)</p> <p>The Human Anatomic Gene Expression Library (H-ANGEL), the H-Inv Integrative Display of Human Gene Expression across Disparate Technologies and Platforms. Tanino M, Debily MA, Tamura T, Hishiki T, Ogasawara O, Murakawa K, Kawamoto S, Itoh K, Watanabe S, de Souza SJ, et al. Nucleic Acids Research 33(Database Issue): D567-D572. (2005)</p>
達成度評価の方法	テストによる評価
備考	

科目名称	プロテオーム解析におけるデータ解釈 (総合判断・産業応用)
区分	必修
講義形態	座学と実習
時間	3時間 (1コマ)
対象者	カリキュラム上級コース受講者 Bio系分野 / IT系分野所属 大学博士課程以上で分子生物学・生命科学を学んだ者
対象知識項目	プロテオーム解析のための総合判断・産業応用の知識
達成目標	プロテオーム解析において解析結果の解釈に必要な機能推定、複数の解析方法による結果の比較・検討、機能の応用利用について習得する。
科目概要	<p>1 解析結果の解釈</p> <p>1.1 質量誤差の評価</p> <p>1.2 翻訳後修飾の同定</p> <p>1.3 タンパク質間相互作用</p> <p>1.4 遺伝子ネットワーク解析</p> <p>1.4.1 プーリアンモデル</p> <p>1.4.2 ペイジアンモデル</p> <p>1.4.3 S-system</p> <p>1.5 システムバイオロジー</p> <p>1.6 総合的解釈</p> <p>1.6.1 機能推定</p> <p>1.6.2 他の解析法との比較</p> <p>1.6.3 疾患等との関連</p> <p>1.6.3.1 疾患原因遺伝子</p> <p>1.6.3.2 診断マーカー</p> <p>1.6.4 タンパク質機能の産業応用</p>
参考文献・資料	<p>実験医学別冊 決定版プロテオーム解析マニュアル 磯辺俊明, 高橋信弘 (編) 羊土社 (2004)</p> <p>実験医学別冊 タンパク質実験ハンドブック 分離・精製, 質量分析, 抗体作製, 分子間相互作用解析などの基本原理と最新プロトコール 総集編</p>

	竹縄 忠臣 (編) 羊土社 (2003)
達成度評価の方法	テストによる評価
備考	

科目名称	タンパク質間ネットワーク解析におけるデータ解釈 (総合判断・産業応用)
区分	必修
講義形態	座学と実習
時間	3 時間 ( 1 コマ )
対象者	カリキュラム上級コース受講者 Bio系分野 / IT系分野所属 大学博士課程以上で分子生物学・生命科学を学んだ者
対象知識項目	タンパク質間ネットワーク解析のための総合判断・産業応用の知識
達成目標	タンパク質間ネットワーク解析において解析結果の解釈に必要な機能推定、複数の解析方法による結果の比較・検討、機能の応用利用について習得する。
科目概要	1 解析結果の解釈 <ul style="list-style-type: none"> <li>1.1 質量誤差の評価</li> <li>1.2 翻訳後修飾の同定</li> <li>1.3 タンパク質間相互作用</li> <li>1.4 遺伝子ネットワーク解析 <ul style="list-style-type: none"> <li>1.4.1 ブーリアンモデル</li> <li>1.4.2 ベイジアンモデル</li> <li>1.4.3 S-system</li> </ul> </li> <li>1.5 システムバイオロジー</li> <li>1.6 総合的解釈 <ul style="list-style-type: none"> <li>1.6.1 機能推定</li> <li>1.6.2 他の解析法との比較</li> <li>1.6.3 疾患等との関連 <ul style="list-style-type: none"> <li>1.6.3.1 疾患原因遺伝子</li> <li>1.6.3.2 診断マーカー</li> </ul> </li> <li>1.6.4 タンパク質機能の産業応用</li> </ul> </li> </ul>
参考文献・資料	実験医学別冊 決定版プロテオーム解析マニュアル 磯辺俊明, 高橋信弘 (編) 羊土社 (2004)  実験医学別冊 タンパク質実験ハンドブック 分離・精製, 質量分析, 抗体作製, 分子間相互作用解析などの基本原理と最新プロトコール 総集編 竹縄 忠臣 (編) 羊土社 (2003)
達成度評価の方法	テストによる評価
備考	

## 第6章 スキルスタンダード・カリキュラムの活用について

### 6 1. スキルスタンダード・カリキュラムの活用方法

本事業におけるスキルスタンダード・カリキュラムは、前述したように、以下の人材を育成対象として想定し、作成された。

- 1) 専任アノテーターという高度バイオ人材の育成  
最終目標であるトップレベルの人材であり、上級レベルとして想定した。
- 2) 企業内・研究所内においてアノテーションをおこなう人材（兼任アノテーター）の育成  
社内データベースを活用する企業でのアノテーション業務を遂行する人材を初級～中級レベルとして想定した。
- 3) 生物学・生命科学系人材が必要とするバイオインフォマティクス基本技能の習得  
アノテーション技能の習得が、実験から産出される生物データを解析、解釈するバイオインフォマティクス技術を含んでいるので、その基本技能の習得を初級～中級レベルとして想定した。

また、本スキルスタンダードは作成時に、利用しやすいように、5つの実験分野ごとに作成しているので、例えばタンパク質関連のアノテーション業務に特化して研修するなど、各ユニット別での活用が可能である。したがって、活用方法として以下の幅広い活用がなされることを期待している。

#### (1) 企業・団体別活用法

- 1) 大規模データベース作成・利用機関、センター  
『人事（採用、評価）・育成』
  - ・アノテーション技能をもつ人材の採用、評価
  - ・専任（上級）アノテーター育成のための参考資料このカテゴリーの研究所の方々には、本スキルスタンダード・カリキュラムへのご意見をいただき、よりの確なスキルスタンダード・カリキュラムへのブラッシュアップを行いたいと考えております。ぜひご意見をお寄せいただきたいと思います。
- 2) 生物学的データ処理を必要とする研究機関、製薬会社、CRO、食品・環境等のバイオ系企業  
【人事（採用、評価）・育成】
  - ・アノテーション技能をもつ人材の採用、評価
  - ・兼任アノテーター、アノテーション能力育成のための参考資料【社内・社外の研修での目標設定】
  - ・アノテーション社内研修の計画時の目安
  - ・アノテーション社外研修の活用とその評価【委託・受託時及び共同研究時の基準】
  - ・企業同士、研究所同士の連携時、発注時の基準
- 3) 医学研究者、バイオ研究者、バイオ系研究所・企業  
【バイオインフォマティクス参入時の参考資料】
  - ・バイオ系研究所がインフォマティクス分野を整備・拡充する際に、フロー・人材面での参考となる資料

- ・バイオインフォマティクス分野に参入する企業の参考資料
- ・バイオ/医療研究者がバイオインフォマティクスに取り組む際の道標

【人事面（採用、評価）】

- ・アノテーション技能をもつ人材の採用、評価

【社内・社外の研修での目標設定】

- ・アノテーション社内研修の計画時の目安
- ・アノテーション社外研修の活用とその評価

【委託・受託時及び共同研究時の基準】

- ・企業同士、研究所同士の連携時、発注時の基準

4) バイオインフォマティクス関連企業

【人事面（採用、評価）】

- ・アノテーション技能をもつ人材の採用、評価

【社内・社外の研修での目標設定】

- ・アノテーション社内研修の計画時の目安
- ・アノテーション社外研修の活用とその評価

5) 教育機関、バイオ系専門学校

【バイオインフォマティクス参入時の参考資料】

- ・バイオ系研究機関、バイオ系専門学校がインフォマティクスに取り組む際に導入しやすい参考資料
- ・バイオ教育者がバイオインフォマティクスに取り組む際の道標

(2) アノテーション技能の概念を広く周知

バイオインフォマティクス関係で世界の動向に精通されている理化学研究所の八尾先生が、最近、日経 BTJ/HEWDLINE/NEWS に以下の趣旨の意見を述べられた。

「2005年1月に、信頼できる公開バイオデータベースのほぼ全貌が分るものとして権威がある Nucleic Acids Research 誌 Database Issue が発行されたが、ここに掲載されているバイオデータベースの数は、2003年386種、2004年548種、2005年719種と急増している。また各々のデータベースに含まれるデータ件数も急増している。

多様なデータベースを有効に活用できるかどうかは、研究者個人、研究グループにとって、大きくは国レベルでも、非常に重要な問題であり、特に現在大きな課題は、各種ゲノムあるいは cDNA 中の未知遺伝子の同定及び既知情報の付与を行ういわゆる「アノテーション」の力である。分かりやすく言えば、最も美味しい実を見つけることができるかどうかがこの業務である。

これまで「アノテーション」は、基本的なゲノム情報などに対し網羅的に注釈付けをするような大規模な動きをして来たが、今後はさらに各研究者・研究グループ・各企業で個々のニーズに応じてアノテーションすることが並行して必要になってくる。

データベースの種類と量は今後ますます増えて行くことは必至で、その「宝庫」から有用な情報を引き出し、研究成果を挙げていくために、「アノテーション人材育成」をさらに進めていく必要を痛感する。」

本スキルスタンダード及びカリキュラムは、アノテーションの業務内容を具体的に記載しており、これらが活用されることによって、第一に「アノテーション」という言葉、業務概念が広まり、その重要性が認識されることが期待される。

#### 1) アノテーターという職業の認知

今後ますます生命科学系のデータベースが増加することより、アノテーターは、生命科学の研究者・ポスドク・修士の人材にとって、目標とすべきキャリアの一つになりうる。信頼できる生命科学系のデータベースの増加が日本でより起こることを希求するとともに、アノテーターという専任職が本スキルスタンダードの公開により、認知が高まることが期待される。

#### 2) アノテーションの示す技術概念の認知

アノテーションは、重要な技能であるが、研究者によってその言葉のイメージが異なる。本アノテーション・スキルスタンダードでは、実験・そのデータの解析・データの解釈・データベースへの格納というアノテーションの流れに沿ってスキルスタンダードを構成しており、この一連の作業ができる人材及び技能として捕らえることができる。本事業が必ずしも唯一無二の考え方ではなく、研究者の言葉の概念を形成するためのたたき台となり、この重要な技能をよりの確に表現する言葉、またはアノテーションという言葉が人口に膾炙することが期待される。

## 6 2 . 次年度以降の展開方針

本事業の成果を踏まえ、以下の事業を次年度に計画している。

### ( 1 ) アノテーションスキル講習会・セミナーの実施

作成したアノテーション・スキルスタンダード及びカリキュラムをベースに講習会・セミナーの開催を予定している。対象は、企業の人材、医療関係者を想定しているので、長期間のコース設定をおこなうのではなく、テーマを絞った1回または4回連続程度のセミナーを、定期的におこなう予定である。

#### 【発現プロフィール(マイクロアレイ)・アノテーション講座】(講習会例)

対象：生物学・生命学を修了した企業、研究所、医療関係者で、マイクロアレイの実験データを持ち、発現プロフィールからアノテーションを行う予定のある人材。企業の中途研修、新人研修。

実施期間： 毎週1回×4回

実施時期： 5月中旬～6月中旬を予定

また、講習会の実施については、一企業で行うよりも、今年度バイオ人材育成システム開発事業で再委託を受けた事業者働きかけ、複数の事業者で行えるように企画したいと考えている。バイオ人材育成の講習会は、過去の経験から事業ベースに乗りやすく、事業者が集まることにより、広報面でのインパクト、経費面での節減を図り、またバイオ技術研修を希望する企業・団体・個人にとっても選びやすい場を提供できる。これによって従来単一の講座では参加しなかった領域・職種を持つ人材が新たに発掘されるだけでなく、各講座を縦断する境界領域も理解できる人材の教育も期待される。

### ( 2 ) バイオインフォマティクス技術認定制度との連携

2004年10月より、社団法人バイオ産業情報化コンソーシアム(JBIC)の主催で、「バイオインフォマティクス技術検定制度」が始まっている。この検定制度は、2003年度JBICの下で、バイオインフォデザイン・ジャパン株式会社、株式会社国際バイオインフォマティクス研究所、株式会社マホレックスが受託しておこなった「バイオインフォマティクス人材育成システム開発事業」を基として、今年度JBICにより、自主事業として開発・発展させたものである。

専任アノテーターは、広義にはバイオインフォマティクスの上級人材の位置付けができ、人材の具体的な就業先であるので、JBICの行う検定制度の拡充を待って、その中に検定制度として位置づけることも検討に値すると思われる。

## 参考資料

### A. 「バイオ産業を支える人材の育成に関する調査（アノテーター）」結果

アノテーション関連事業・研究の実施状況及び人材に関する標記調査を行なったところ、37件の回答を得た。

(3) 貴社もしくは貴研究所では、アノテーション関連事業・研究に取り組んでいますか。

はい	いいえ	無回答
24	11	2
(65%)	(30%)	(5%)

(4) 「アノテーター」について、貴社もしくは貴研究所にとっての理想的な人材像・人材イメージをご自由にご記入ください。

- ・ 一般的な生物学の知識をもち、更にこの分野における研究の最新情報を理解している、もしくは理解する姿勢をもっている。またコンピューターに関する知識・能力を持っている必要がある。特にアノテーションされた情報の管理・更新・公開などでは研究者を補佐しなくてはいけないため、データベース管理能力は特に必要である。
- ・ アノテーションの対象がどういったものであるのかを理解すること
- ・ アノテーション付与の意味を正確に理解すること
- ・ データ内容・構成・分類などの構造的なことを理解し、データをまとめるデータを蓄積する方法論を提案できる
- ・ 営利目的ではないため、当方での実状はあまり参考にならないのではと考える。強いて想像するならば、研究もしくは事業の目的を見失わないこと、合目的性、要領の良さのような「嗅覚」が必要とされそうに思う。
- ・ 個々の専門分野に精通して、関連する情報を幅広く収集でき、その分野のスペシャリストとなれる複数の人材。
- ・ Biological な知識 + Database Administration
- ・ コンピューターの操作、検索技術が問題では勿論のこと、しなやかさのある思考力をもった人材の育成が重要。生物学を良く理解している人材、研究課題、個別に興味があることもスタートとして大切であるが、全体を俯瞰的にとらえることができる人材・興味の範囲が広い人材育成。
- ・ 生化学、分子生物学の広範な実験手法を実施したことがある人
- ・ 学術論文を年間 500 報以上目を通す人
- ・ 形式でなく、内面を理解している。
- ・ 細胞だけでは、出口である医学へつながらないことを理解している。
- ・ 生物研究の意味、目的、出口を理解していなければいけない。
- ・ 調査票に書かれている定義と全く同じ(他の研究者やデータベースシステム開発者と共同して業務を進める必要から特定の専門分野に限らず幅広い知識とコミュニケーション能力が必要とされ、アノテーターがこれらの構成員の中で橋渡し役として中心的な役割を果たすことが期待されている)
- ・ バイオインフォマティクス解析技術やデータベース構築技術は、微生物のゲノム解析に基づく病原性遺伝子の探索はもちろんのこと、有用微生物における食品産業や健康食品の開発にも多くの需要があることが近年分かってまいりました。しかし、こうした研究をサポートできるアノテーターの人数は少なく、むしろ需要の拡大を抑え、必要とされているニーズに対応できないのが現状です。そこでアノテーターに求められるものとしましては、遺伝情報や遺伝的背景を正しく理解でき、適切なプライバシー保護に基づいた健康管理あるいは健康食品などを推進し、多くの人々

が求めているニーズに応えることこそが、最も望まれる人材像であると考えています。

- ・ バイオインフォマティクスの有効性、活用方法を理解し、活用していただける人材
- ・ 分子生物学と情報学の知識、経験を兼ね備えた人材
- ・ アノテーションの重要性・影響力の大きさを強く認識していること。
- ・ タンパク質についてのイメージを持っている人。  
偏見を拭い去ることができる人。
- ・ 分子生物の知識をもち、ツールやデータベースを使いこなせる or 論文の内容を要約できる IT 技術者
- ・ 特定の分野に特化するのではなく、生物学、情報学といった幅広い知識を有する方
- ・ アノテーションデータからの新規情報の発掘を行なえる方
- ・ アノテーションに関する作業を効率化のためのシステム開発を行なえる方
- ・ 地道な努力家。幅広い知識あるいは知識獲得に積極性のある人間。
- ・ 生物とバイオインフォマティクスの両者を理解でき、アウトプットが出せる
- ・ 創薬と分子生物学の基礎的知識があり、適切な用語で機械処理可能な記述ができる、オントロジーや自然言語処理のシステムにも詳しい人材。
- ・ 生物学の発展の最先端を常に意識し把握されること。
- ・ 国際的に活動ができること。
- ・ GPCR、キナーゼなどのファミリーごとに、配列情報や機能情報、構造情報、特許情報など様々な情報を整理・統合でき、更新維持できる人。各種公開DBを適確に選別し、統合できる人。未知遺伝子の機能推定するツールを一通り知っていて、使える人。

(5) アノテーター人材に望まれる能力は

研究者・科学者としての思考がある	17	(46%)
専門領域をもっている	13	(35%)
最新技術に精通している	12	(32%)
経験が豊富である	10	(27%)
技術力が高い	9	(24%)
コミュニケーション・交渉能力がある	9	(24%)
プロジェクト管理能力が高い	8	(22%)
人脈が豊かで社内外の連携体制がとれる	5	(14%)
市場・顧客ニーズを把握できる	3	(8%)
リーダーシップに優れている	1	(3%)
その他	2	(5%)
特にない	0	0%

(6) アノテーター人材像のレベル分類仮説について

	概ね妥当	妥当ではない
<b>初級者：</b> アノテーション・スタンダード	23 (92.0%)	2 (8.0%)
<b>中級者：</b> アノテーション・エキスパート	22 (91.7%)	2 (8.3%)
<b>上級者：</b> アノテーション・マネージャー	22 (91.7%)	2 (8.3%)

[妥当ではない場合、必要な能力・知識、役割など]

- (初級) 具体的な技術手法 (SQL, XML Parsing 等) の理解があるべきである。
- (中級) コミュニケーションできることは能力ではない。自分のアイデアを提言できるなどの能力を持つことが必要。
- (上級) バイオインフォマティクス 実験者だけでは不十分。中心にはなれないと思われる。生物研究の出口とは何かを再考すべき。

(7) アノテーターに必要なスキル等について

分野	項目	極めて必要	必要	あまり必要ではない	不要	最頻値	
配列解析(シーケンス)	実験の知識・経験	実験目的、実験原理の理解	6	16	4	0	2
		実験装置、実験方法の理解	5	14	6	1	2
	検索・予測ツールの知識・経験	検索データベースの知識	11	14	1	0	2
		検索・予測ツールの原理の理解	12	14	0	0	2
		検索・予測ツールの活用	9	15	2	0	2
	検索・予測結果の解釈	分子生物学の知識	13	12	1	0	1
		統計学の知識	7	16	3	0	2
		総合判断・産業応用	10	10	5	1	1
	発現プロファイル解析(DNAマイクロアレイ・DNAチップ)	実験の知識・経験	実験目的、実験原理の理解	2	17	6	1
実験装置、実験方法の理解			2	14	8	2	2
検索・予測ツールの知識・経験		解析ツールの知識	6	12	5	3	2
		解析ツールの原理の理解	5	12	7	2	2
		解析ツールの活用	5	14	5	2	2
検索・予測結果の解釈		分子生物学の知識	11	10	4	1	1
		統計学の知識	6	15	4	1	2
		総合判断・産業応用	9	10	4	3	2
多型解析(SNP)		実験の知識・経験	実験目的、実験原理の理解	4	11	8	3
	実験装置、実験方法の理解		3	12	8	3	2
	検索・予測ツールの知識・経験	検索データベースの知識	4	17	4	1	2
		検索・予測ツールの原理の理解	5	15	4	2	2
		検索・予測ツールの活用	4	16	4	2	2
	検索・予測結果の解釈	分子生物学の知識	7	15	3	1	2
		統計学の知識	6	15	4	1	2
		総合判断・産業応用	5	14	5	2	2

プロテオーム解析(質量分析)	実験の知識・経験	実験目的、実験原理の理解	4	14	6	2	2
		実験装置、実験方法の理解	3	11	10	2	2
	検索・予測ツールの知識・経験	検索データベースの知識	7	16	3	0	2
		検索・予測ツールの原理の理解	2	17	5	2	2
		検索・予測ツールの活用	3	15	6	2	2
	検索・予測結果の解釈	分子生物学の知識	8	12	6	0	2
		統計学の知識	2	17	6	1	2
		総合判断・産業応用	6	12	7	1	2
	タンパク質間ネットワーク(遺伝子ネットワーク)解析	実験の知識・経験	実験目的、実験原理の理解	4	15	6	1
実験装置、実験方法の理解			4	13	8	1	2
検索・予測ツールの知識・経験		検索データベースの知識	6	14	6	0	2
		検索・予測ツールの原理の理解	2	15	7	2	2
		検索・予測ツールの活用	3	15	5	3	2
検索・予測結果の解釈		分子生物学の知識	5	15	6	0	2
		統計学の知識	2	16	7	1	2
		総合判断・産業応用	6	13	6	1	2
資質・志向		国際性	データベースの国際的な基準への準拠	9	15	1	1
	信頼性	データの信頼度を高める一貫性、継続性、確実性	11	14	1	0	2
	前進意欲	生命科学の急速な進展と並走する前進性	12	11	3	0	1
	倫理性	実務の倫理性	7	16	3	0	2

アナレーターに必要とされるスキル、特に配列解析に関して必要を強く訴える回答が多く見られた。これは、他の分野においても基本的に必要とされる「必修科目」であることを示唆していると思われる。また、どの分野も「極めて必要」「必要」により解答が集中している。特にデータベース知識と分子生物学の知識は、どの分野でも「極めて必要」とした回答が多かった。

## (8) スキルの不足感について

分野	項目		極めて 不足している	やや不足している	あまり不足していない	不足していない	最値 頻
配列解析 (シーケンス)	実験の知識・ 経験	実験目的、実験原理の理解	0	8	9	8	3
		実験装置、実験方法の理解	0	8	10	7	3
	検索・予測ツールの知識・ 経験	検索データベースの知識	2	9	11	2	3
		検索・予測ツールの原理の理解	2	11	8	3	2
		検索・予測ツールの活用	4	12	6	2	2
	検索・予測結果の解釈	分子生物学の知識	0	15	5	4	2
		統計学の知識	0	14	7	3	2
		総合判断・産業応用	2	11	6	5	2
	発現プロファイル解析 (DNAチップ)	実験の知識・ 経験	実験目的、実験原理の理解	0	11	10	4
実験装置、実験方法の理解			0	12	9	4	2
検索・予測ツールの知識・ 経験		解析ツールの知識	0	9	10	5	3
		解析ツールの原理の理解	0	9	11	4	3
		解析ツールの活用	0	9	11	4	3
検索・予測結果の解釈		分子生物学の知識	1	13	5	5	2
		統計学の知識	1	14	6	3	2
		総合判断・産業応用	2	12	4	6	2
多型解析 (SNP)		実験の知識・ 経験	実験目的、実験原理の理解	2	7	11	4
	実験装置、実験方法の理解		2	7	9	5	3
	検索・予測ツールの知識・ 経験	検索データベースの知識	1	11	8	4	2
		検索・予測ツールの原理の理解	1	11	9	3	2
		検索・予測ツールの活用	1	11	9	3	2

	検索・予測結果の解釈	分子生物学の知識	0	11	8	5	2
		統計学の知識	3	10	5	6	2
		総合判断・産業応用	2	11	5	6	2
プロテオーム解析（質量分析）	実験の知識・経験	実験目的、実験原理の理解	3	12	6	4	2
		実験装置、実験方法の理解	2	8	10	4	3
	検索・予測ツールの知識・経験	検索データベースの知識	2	11	7	4	2
		検索・予測ツールの原理の理解	3	10	6	5	2
		検索・予測ツールの活用	3	10	6	5	2
	検索・予測結果の解釈	分子生物学の知識	3	10	5	5	2
		統計学の知識	2	10	6	5	2
		総合判断・産業応用	2	11	5	5	2
	タンパク質間ネットワーク（遺伝子ネットワーク）解析	実験の知識・経験	実験目的、実験原理の理解	1	14	5	4
実験装置、実験方法の理解			1	11	8	4	2
検索・予測ツールの知識・経験		検索データベースの知識	0	14	6	4	2
		検索・予測ツールの原理の理解	2	11	6	5	2
		検索・予測ツールの活用	2	11	5	6	2
検索・予測結果の解釈		分子生物学の知識	0	12	6	5	2
		統計学の知識	1	11	6	5	2
		総合判断・産業応用	3	10	6	4	2
資質・志向		国際性	データベースの国際的な基準への準拠	1	11	8	4
	信頼性	データの信頼度を高める一貫性、継続性、確実性	1	13	7	3	2
	前進意欲	生命科学の急速な進展と並走する前進性	1	9	10	4	3
	倫理性	実務の倫理性	3	9	7	4	2

スキルの不足感に関しては、強い不足感は認められなかったものの、ほとんどの解答について中庸よりむしろ「やや不足している」に傾いている。

(9) 各スキルの育成・獲得方法について

分野	項目	社内研修	社外研修	OJT	育 成 し て い い	最 値 頻	
配列解析 (シーケンス)	実験の知識・ 経験	実験目的、実験原理の理解	7	1	9	7	3
		実験装置、実験方法の理解	7	1	8	7	3
	検索・予測ツールの知識・ 経験	検索データベースの知識	8	1	8	5	1,3
		検索・予測ツールの原理の理解	8	2	8	4	1,3
		検索・予測ツールの活用	8	1	8	5	1,3
	検索・予測結果の解釈	分子生物学の知識	8	3	7	3	1
		統計学の知識	8	2	7	4	1
		総合判断・産業応用	6	2	8	6	3
	発現プロファイル解析 (DNAマイクロアレイ・DNAチップ)	実験の知識・ 経験	実験目的、実験原理の理解	7	3	6	6
実験装置、実験方法の理解			7	3	6	6	1
検索・予測ツールの知識・ 経験		解析ツールの知識	7	3	5	7	1,4
		解析ツールの原理の理解	7	3	5	7	1,4
		解析ツールの活用	7	3	6	6	1
検索・予測結果の解釈		分子生物学の知識	6	3	6	5	1,3
		統計学の知識	7	3	5	6	1
		総合判断・産業応用	5	3	4	9	4
多型解析 (SNP)		実験の知識・ 経験	実験目的、実験原理の理解	4	3	6	9
	実験装置、実験方法の理解		3	3	7	8	4
	検索・予測ツールの知識・ 経験	検索データベースの知識	8	3	5	6	1
		検索・予測ツールの原理の理解	8	2	7	5	1
		検索・予測ツールの活用	8	3	6	5	1

	検索・予測結果の解釈	分子生物学の知識	6	4	6	5	1,3
		統計学の知識	6	4	4	7	4
		総合判断・産業応用	5	3	5	8	4
プロテオーム解析（質量分析）	実験の知識・経験	実験目的、実験原理の理解	5	4	5	7	4
		実験装置、実験方法の理解	5	2	6	8	4
	検索・予測ツールの知識・経験	検索データベースの知識	7	2	5	7	1,4
		検索・予測ツールの原理の理解	6	3	5	7	4
		検索・予測ツールの活用	6	3	5	7	4
	検索・予測結果の解釈	分子生物学の知識	6	3	6	6	1,3,4
		統計学の知識	7	3	6	6	1
		総合判断・産業応用	6	2	6	8	4
	タンパク質間ネットワーク（遺伝子ネットワーク）解析	実験の知識・経験	実験目的、実験原理の理解	5	3	5	8
実験装置、実験方法の理解			5	3	5	8	4
検索・予測ツールの知識・経験		検索データベースの知識	7	3	6	6	1
		検索・予測ツールの原理の理解	6	4	5	7	4
		検索・予測ツールの活用	6	4	4	8	4
検索・予測結果の解釈		分子生物学の知識	6	4	5	6	1,4
		統計学の知識	7	3	4	7	1,4
		総合判断・産業応用	6	3	5	7	4
資質・志向		国際性	データベースの国際的な基準への準拠	7	2	5	7
	信頼性	データの信頼度を高める一貫性、継続性、確実性	6	2	8	5	3
	前進意欲	生命科学の急速な進展と並走する前進性	8	2	7	4	1
	倫理性	実務の倫理性	7	4	3	6	1

各スキルを獲得する方法としては、社内教育、On the Job Training というケースが多い割に、社外研修が行われるケースは著しく少ない。組織内では必要性が認められ

ている割に、社外でこうしたスキル獲得のための研修を受ける機会を得にくいことが読み取れる。

(10) 挙げられているスキル項目以外にも重要と考えられるスキル項目は

- ・ 複合グラフモデル、パスウェイ記述法(XML、ペトリネット等)
- ・ タンパク質・アミノ酸についての物理的性質の知識を持つこと
- ・ テキストマイニング、原理と利用
- ・ 頭の中に物体をイメージし、相互作用をイメージの中で操作できること

(11) 現在の生物データベースやアノテーションにおいて不備な点や改善が望まれる点は

ある	特にない	わからない	無回答
15	2	3	17
(40.5%)	(5.4%)	(8.1%)	(45.9%)

[不備な点]

- ・ アノテーションされた内容の信頼性の評価。具体的な例を挙げると遺伝子配列やタンパク質配列の機能に関して、実際に実験で確かめられたものなのか、予測結果なのか記載に統一性がないためにデータの解釈が困難である。
- ・ これはデータベースというよりも一般の研究において語彙の統制が不足していたことと、現状 GO などの試みも未完のため、語彙を制御しない、もしくは、これを軽視する傾向が強い、公共インフラがない場合でも独自のもを必要なら作成するほうが望ましい。
- ・ 多種多量の情報が混在する中で、それらの情報の信頼性や基準がバラバラである。データフォーマットの統一だけでなく、それら情報の根拠、信頼性がある程度把握できるような整備が望まれると考えます。
- ・ データフォーマットの検討。
- ・ ユーザフレンドリーにすること。
- ・ 実験で生物学的キノウが立証されたものと配列の相同性からキノウを推定したものの違いがわかりにくい。
- ・ Ontology の充実。
- ・ フィロソフィーが重要。生物DB研究の出口をにらんだ形でのフィロソフィー設定がなければならない。
- ・ 分野として急速に発展したため、全体的に経験不足。アノテーターが配列解析に偏りすぎ。Wet 研究者との連携不足。
- ・ アノテーションの情報の由来がわからない。
- ・ 詳細な基準が明確ではない(明確に書かれていない)。
- ・ データベースはフラットになりがちだが、生物の階層の見通しのよいデータベースがあると良いと思う。
- ・ 情報の統一化
- ・ 必要情報を可不足なく抽出できる検索機能
- ・ 解析手法の統一化
- ・ 情報の精査方法の確立
- ・ 生物データベースのアノテーション不足 アノテーションの間違いの訂正の不足 生物学の知識変化への対応。
- ・ database の個性も重要かとは思いますが、統一性がない為に、表記法なども多様。Ontology の進化で問題が解消あるいは、少なくなるかもしれないが。
- ・ データの品質(不正規、間違い、信頼性) 2.カバー率 3.オントロジー対応
- ・ 特に発現プロファイリングDBの不備、統一性のないことが、大きな問題と思われる。

- ・ アノテーターの教育が不備
- ・ アノテーションの重要性が未だ社会的に認識されていない。
- ・ 日本が発信源となっている DB は、更新・継続性が不安定なものが多く、定常的に使用することが少ない。

(12) スキル項目に沿ったカリキュラムの外部研修(有料)があった場合、貴社もしくは貴研究所では、従業員、所員を外部研修に参加させたいと思いますか

参加させたい	無料なら参加させてもよい	参加させたいとは思わない	無回答
10	19	2	5
(27.0%)	(51.4%)	(5.4%)	(13.5%)

[金額について]

5万円程度(2)、10万円程度(2)、2~3万円、3~4万円、7~10万円

[日数について]

~2日	4
~4日	7
~6日	8
~8日	3
~10日	4
11日~	1

(13) 本事業の成果を活用してアノテーターの検定試験が実施された場合、貴社もしくは貴研究所では、従業員、所員に検定を受検させたいと思いますか

ぜひ受験させたい	内容によっては受験させてもよい	受験させたいとは思わない	無回答
3	22	6	5
(8.3%)	(61.1%)	(16.7%)	(13.9%)

(14) 業界標準のスキルスタンダードが作成された場合、貴社もしくは貴研究所では、どのような観点から利用できると思いますか

採用活動時における能力評価基準	16	(43.2%)
人事考課時の判断基準	3	(8.1%)
人材配置の判断基準	14	(37.8%)
社員に必要な能力開発の明確化	18	(48.6%)
社内人材ニーズの把握	5	(13.5%)
その他	4	(10.8%)

[その他の意見]

人材育成カリキュラム、学生の就職に有利

(15) 本事業や、広くバイオテクノロジーを支える人材の育成に対するご意見・ご要望等

- ・ 人材の育成ばかりではなく、人材を活用するための環境整備も進めていただきたい。例えば、人材派遣を予算や期間を柔軟に設定できる形で提供する仕組みなどは、大学の研究室を運営するものにとって大変有難いものになる。

- ・ バイオサイエンスの各分野について検定試験があり、就職活動が有利になるのであれば、学生に積極的に受験をすすめたい。
- ・ 生命現象に興味を持ち、その解明と応用に使命感をもった人材が多くなるのが最も重要。  
そのためにはあらゆる分野の人間が（生物系のみならず、物質系・情報系の人も）対等に活躍できるチャンスがあることを示し、他分野の集団（学会、業界）に参画を呼びかけることが望ましい。
- ・ バイオテクノロジーは大きく変わっていくと思う。あまり固定的に考えない方が良いと思う。
- ・ 技術者の養成とともに研究者の養成も重要。
- ・ 「アノテーション」と言うこと自身の定義がまず現在は無いのではないか。例えばSNPの解析データから遺伝子の重要データのマイニングを行う事がアノテーションなのか、塩基配列から遺伝子領域を予測することがアノテーションなのか遺伝子の機能予測をすることがアノテーションなのか、ということに基準がまず存在しない。さらに遺伝子の機能予測に関しても、相同性が有るという事で機能予測が出来たとするか、その時の相同性のスレッショールドはどの位の値にするか、その相同性を示す既知遺伝子の情報と照らし合わせるのか。これらの事項についてはスタンダードが無いのが現状で、各研究機関において個別の基準を設けて対応していると思われる、その基準が異なる中ではどのような研究者技術者を「アノテーター」と呼ぶのかも違っていると思われる、そのような中で問5、6のようなレベルを比較しても余り意味が無いのではないか？「アノテーター」の育成か検定試験を行うよりは、バイオの知識とコンピューターの知識を同時に学べる学際的な教育機関を設ける（既に欧米では進んでいる）ことが有意義だと思われる。
- ・ アノテーターの人材育成は大変重要であると思います。また育成した人材が能力を存分に発揮できる職場が必要であり現在は後者の確保がより急務と感じます。しかも安定した職場であることを望みます。日本人は几帳面で多方面の潜在能力が高いので、優れたアノテーターが職業のひとつとして認められることを期待いたします。独立した職業として。
- ・ 生物学データを情報科学、統計学の知識・手法を用いて解析・解釈できるような人材を育てる教育機関が不足していると感じているので、非常に良い試みと思う。また数日のコースとしてだけでなく、トピック毎に1日程度の研修もあると良い
- ・ We t作業と両立しないといけないと思います。
- ・ 人材育成が極めて重要である。
- ・ このアンケートがまだ実施されていなかったとは夢にも思わなかった。
- ・ アノテーターは、技術者なのか研究者なのか。技術者であれば、国家的プロジェクトのようなどころ以外に、どこで仕事ができるのか。人生設計として一生の仕事になり得るのか。研究者であればアノテーションは、どれ程主体的研究課題たり得るのか。あり得るひとつの解答は、他に主務を持つ者が、さらにアノテーターの技能を身につけるといふものだが、片手間でできる技術だろうか。

B. アノテーター実証講座講義記録

(1) 第1回

日時	2004年10月13日
科目名	配列解析1
時間・コマ数	3時間(1コマ)
講師名	大城戸 利久
講義概要	<ol style="list-style-type: none"> <li>1 生命科学における情報とは？</li> <li>2 セントラルドグマ</li> <li>3 遺伝子と産物(タンパク質); 真核生物</li> <li>4 遺伝子と産物(タンパク質); 原核生物</li> <li>5 生物の解析</li> <li>6 配列の解析</li> <li>7 生命科学データベース</li> <li>8 2004年NAR誌(Nucleic Acid Research)のDatabaseの分類</li> <li>9 タンパク質データベース</li> <li>10 タンパク質モチーフ検索</li> <li>11 コンピューターを用いた遺伝子発見の方法</li> <li>12 代表的な遺伝子発見プログラム</li> <li>13 核酸, タンパク質の配列類似度検索</li> <li>14 遺伝子産物の機能分類</li> <li>15 実習内容</li> <li>16 実習1 塩基配列の機能推定</li> <li>17 DDBJ(日本DNAデータバンク)のトップページ</li> <li>18 Nucleotide Sequence Database</li> <li>19 DDBJから取得可能な核酸, タンパク質配列データ</li> <li>20 DDBJから取得可能な解説および参考文献</li> <li>21 DDBJからデータ取得可能なミラーデータベース</li> <li>22 DDBJサービス(データベース検索)</li> <li>23 塩基配列の取得(1)</li> <li>24 getentryのメニュー</li> <li>25 塩基配列の取得(2)</li> <li>26 タンパク質コード領域の推定(1-1)</li> <li>27 タンパク質コード領域の推定(1-2)</li> <li>28 タンパク質コード領域の推定(2-1)</li> <li>29 タンパク質コード領域の推定(2-2)</li> <li>30 タンパク質コード領域の推定(2-3)</li> <li>31 タンパク質コード領域の推定(3-1)</li> <li>32 タンパク質コード領域の推定(3-2)</li> <li>33 cDNA上のタンパク質コード領域の解析結果</li> <li>34 配列相同性検索</li> <li>35 BLAST(Basic Local Alignment Search Tool)プログラム</li> <li>36 FASTAプログラム</li> <li>37 配列相同性検索(BLASTトップページ)</li> <li>38 配列相同性検索(BLASTトップページ)</li> <li>39 配列相同性検索(BLASTトップページ)</li> <li>40 配列相同性検索(検索の実行)配列相同性検索(検索の実行)タンパク質モチーフデータベース</li> <li>41 n t e r P r oデータベースの構成</li> <li>42 タンパク質モチーフ検索</li> <li>43 タンパク質モチーフ検索(結果)</li> <li>44 タンパク質 膜貫通領域検索</li> <li>45 タンパク質 膜貫通領域検索(結果1)</li> <li>46 タンパク質 膜貫通領域検索(結果2)</li> <li>47 タンパク質局在シグナル配列検索</li> </ol>

48	タンパク質局在シグナル配列検索 (結果)
49	タンパク質分泌シグナル配列検索
50	タンパク質分泌シグナル配列検索 (結果)
51	ゲノム上の位置の推定 (ア) ゲノム上の位置の推定 (1) (イ) ゲノム上の位置の推定 (2) (ウ) ゲノム上の位置の推定 (3) (エ) ゲノム上の位置の推定 (4)
52	結果のまとめ
53	実習 2 キーワード検索による情報の抽出
54	データベース検索における関連事項 (1)
55	データベース検索における関連事項 (2)
56	データの一例 (DDBJ のデータ)
57	タンパク質データベース (SWISS-PROT) のエントリーの例
58	検索語
59	検索実習
60	検索結果
61	データ閲覧 (塩基配列)
62	データ閲覧 (OMIM)
63	データ閲覧 (OMIM_ゲノム上の位置_細胞遺伝学的)
64	データ閲覧 (Homologene)
65	データ閲覧 (OMIM_Homologene) BRCA1
66	データ閲覧 (発現情報_GEO2)
67	データ閲覧 (発現情報_GEO2)
68	データ閲覧 (発現情報_GEO3)
69	データ閲覧 (発現情報_GEO5)
70	データ閲覧 (site search2)
71	データ閲覧 (3D structure2)
72	データ閲覧 (3D structure4)
73	データ閲覧 (3D structure_Domain1)
74	データ閲覧 (3D structure4)
75	他の代表的な統合検索サイト

## ( 2 ) 第 2 回

日時	2004 年 10 月 20 日
科目名	配列解析 2
時間・コマ数	3 時間 ( 1 コマ )
講師名	小菅 武英
講義概要	<ol style="list-style-type: none"> <li>1 実習内容</li> <li>2 遺伝子探索の概略(例)</li> <li>3 ホモロジー検索による遺伝子予測</li> <li>4 BLAST プログラム</li> <li>5 BLAST family</li> <li>6 FASTA</li> <li>7 Q1: <i>Salmonella typhimurium</i> (バクテリア)のゲノムからショットガン法により以下の配列が得られた。この中にはどの遺伝子が存在するか。</li> <li>8 BLAST で遺伝子を検索</li> <li>9 blastx 設定画面(1)</li> <li>10 参照 database (protein)</li> <li>11 BLAST 設定画面(2)</li> <li>12 Filter について</li> <li>13 BLAST の検索ステップ</li> <li>14 BLAST 設定画面(2)</li> <li>15 MATRIX</li> <li>16 補足 : MATRIX table</li> <li>17 NCBI BLAST 結果表示</li> <li>18 結果の評価</li> <li>19 手元のコンピュータで BLAST を行う</li> <li>20 Standalone タイプ・www サーバタイプ</li> <li>21 参照データ (FASTA 形式)</li> <li>22 Standalone blast の手順(Windows 版の場合)</li> <li>23 Q2: Standalone blast を利用する</li> <li>24 Q2 の手順</li> <li>25 Q2 の blast 出力ファイル</li> <li>26 Q2 の blast 出力ファイル(alignment 1)</li> <li>27 Q2 の blast 出力ファイル(alignment 2)</li> <li>28 Q2 対象データのチェック 1</li> <li>29 Q2 対象データのチェック 2</li> <li>30 UniProt</li> <li>31 Q9ZG98 のデータチェック</li> <li>32 EC 番号から酵素データを閲覧</li> <li>33 対象データの信頼度</li> <li>34 Answer2</li> <li>35 Q2 の結果を閲覧する</li> <li>36 Viewer の画面</li> <li>37 Q3: <i>Escherichia coli</i> CFT073 株ゲノムには、単独で以下のような配列が存在した。Swiss-Prot のデータを利用して遺伝子コード領域を推定せよ。(ファイルは sequence フォルダ内の sequence03. txt)</li> <li>38 Q3</li> <li>39 Q3 blastx の結果 1</li> <li>40 RF-2</li> <li>40.1 Peptide Chain Release Factor 2</li> <li>41 Q3 blastx 結果 2</li> <li>42 Answer3</li> <li>43 Q4: sequence04. txt 配列は、超好熱性古細菌 <i>Thermococcus</i> sp. のゲノムの一部をしめす。Swiss-Prot database との比較から、どのような遺伝子が存在すると考えられるか</li> <li>44 Answer4</li> <li>45 Q4 blastx(web 版の実行結果)</li> <li>46 ORF の候補を確認</li> </ol>

47	Q4 gene 1
48	ORF の開始、終始コドンを確認
49	Q4 gene 2
50	Q4 gene 3
51	Q4 gene 4
52	Q4 gene 5
53	Q4 gene 6
54	Q4 gene 7
55	Tryptophan 生合成経路と比較
56	Q5: Swiss-Prot を参照データベースに用いて各アミノ酸配列の遺伝子産物を推定せよ。
57	Answer 5 (Swiss-Prot を参照にした場合のプロダクト名)
58	H-Inv の結果と比較
59	2. 比較ゲノム(comparative genomics)
60	Q6: 2 種類の <i>Bacillus licheniformis</i> のゲノムが次のように決定された。 <i>Bacillus licheniformis</i> strain A 株から Blic_strA.txt、 <i>Bacillus licheniformis</i> strain B 株から Blic_strB.txt。この 2 種類についてゲノムレベルでの相違を導き出せ。
61	megablast
62	megablast の出力結果
63	同じ内容を blast で行う場合
64	結果の閲覧
65	Q6-Act.txt の結果を Viewer で閲覧
66	strA, strB 株のゲノムを比較
67	Answer 6
68	Q7: 現在、国際塩基配列データベースには 2 種類の <i>Corynebacterium glutamicum</i> ATCC 13032 株が登録されている。Kyowa 由来 (Kyowa 株とする)、Universitaet Bielefeld 由来 (Bielefeld 株とする)、これら 2 株のゲノムの差異は何か?
69	Answer 7-1
70	Answer 7-2
71	Answer 7-3
72	Answer 7-4
73	3. 遺伝子予測プログラムを使用する
74	遺伝子の予測
75	遺伝子予測プログラム
76	FGENESB (web サーバータイプ)による遺伝子予測
77	Q8: 遺伝子予測プログラムを用いて次のゲノムから coding 領域の候補を導き出せ
78	Answer 8 出力結果

(3) 第3回

日時	2004年10月27日
科目名	配列解析3
時間・コマ数	3時間(1コマ)
講師名	田中 剛
講義概要	<ol style="list-style-type: none"><li>1 実習内容</li><li>2 これまでの実習内容</li><li>3 DNA 配列、アミノ酸配列は世界中で共有されている</li><li>4 配列を比較することで遺伝子の進化過程を推測することができる(アノテーションの応用)</li><li>5 ホモログ(オルソログとパラログ)</li><li>6 系統樹作成法の色々</li><li>7 平均距離法(UPGMA)</li><li>8 最大節約法(Maximum Parsimony)</li><li>9 最尤法(Maximum Likelihood)</li><li>10 N-J法(Neighbor Joining)</li><li>11 CLUSTALX(W)とMEGA2</li><li>12 CLUSTALX</li><li>13 MEGA2</li><li>14 解析1<ol style="list-style-type: none"><li>14.1 ヒトとマウスのCaspase配列を用いて系統樹を作成する</li></ol></li><li>15 目的</li><li>16 方法</li><li>17 系統樹作成結果(CLUSTALX =&gt; NJprot)</li><li>18 結果</li><li>19 解析2<ol style="list-style-type: none"><li>19.1 プラナリアのEST解析で得られた配列の機能予測を行う。</li></ol></li><li>20 目的</li><li>21 配列データ(その一)</li><li>22 配列データ(その二)</li><li>23 ORF Finder (Open Reading Frame Finder)</li><li>24 ORF Finderによりタンパク質配列を予測する</li><li>25 BLASTプログラムを用いて相同配列を直接探る</li><li>26 BLAST検索(1)</li><li>27 BLAST検索(2)</li><li>28 BLASTX結果</li><li>29 Tax Browser</li><li>30 究極の探索方法</li><li>31 相同検索から分かった二つの遺伝子の関係</li><li>32 Kyoto Encyclopedia of Genes and Genomes (KEGG)</li><li>33 LIGAND(化合物カタログ)</li><li>34 Pyridoxine = ビタミン B6</li><li>35 KEGGデータベースから分かったこと</li><li>36 配列の機能推定のまとめ</li><li>37 IUBMB (酵素番号を取りまとめている)</li><li>38 ここまでやってきて疑問を一つ</li><li>39 ゲノム配列に対して検索をかける</li><li>40 真核生物ゲノムデータベース(一例)</li><li>41 各生物種のゲノム配列、予測されたORFに対して相同検索をかける</li><li>42 相同遺伝子の検索結果</li><li>43 ゲノム上の位置を探る</li><li>44 ゲノム上でのSNZ, SNO 遺伝子の位置</li><li>45 遺伝子進化の考察</li><li>46 まとめ</li><li>47 生物学における階層構造をつなげる研究</li></ol>

## (4) 第4回

回・日時	2004年11月10日
科目名	発現プロファイル解析1
時間・コマ数	3時間(1コマ)
講師名	荻島 創一
講義概要	<ol style="list-style-type: none"> <li>1 講義内容</li> <li>2 背景-計算生物学の動向：個々の理解からシステムとしての理解へ</li> <li>3 -ome データの統合と疾患情報との連携</li> <li>4 Gene Ontology</li> <li>5 O B F - Open Bioinformatics Foundation</li> <li>6 文献情報の自然言語処理によるデータベース自動構築</li> <li>7 セマンティック Web</li> <li>8 医学用語と生物学用語のオントロジーの融合の試み</li> <li>9 Systems biology - Life 's Complexity Pyramid</li> <li>10 システムバイオロジーの動向 <ol style="list-style-type: none"> <li>10.1 アメリカ</li> <li>10.2 ヨーロッパ</li> <li>10.3 日本</li> </ol> </li> <li>11 システムバイオロジーの基本戦略</li> <li>12 遺伝子発現データからの遺伝子調節ネットワークの推定</li> <li>13 代表的な細胞のシミュレーション</li> <li>14 E-Cell</li> <li>15 SBML</li> <li>16 ショウジョウバエの概日リズムのシステム解析</li> <li>17 京大細胞・生体シミュレータプロジェクト <ol style="list-style-type: none"> <li>17.1 FDA 治験に耐えうるシミュレータの開発へ</li> </ol> </li> <li>18 システム生物学の医療・創薬への応用</li> <li>19 Pharmacogenetics/genomics との連携</li> <li>20.1 パーソナライズド・メディスンへ</li> <li>20 マイクロアレイによるトランスクリプトーム解析</li> <li>21 網羅的な遺伝子発現解析</li> <li>22 マイクロアレイ解析の例(1)</li> <li>23 マイクロアレイ解析の例(2)</li> <li>24 マイクロアレイバブルの終焉 <ol style="list-style-type: none"> <li>24.1 再現性の高い、成熟したマイクロアレイ解析へ</li> </ol> </li> <li>25 マイクロアレイ解析におけるデータマイニングの特徴 <ol style="list-style-type: none"> <li>25.1 従来の典型的な臨床研究との決定的な違い</li> </ol> </li> <li>26 マイクロアレイ解析をいかに進めるか？ <ol style="list-style-type: none"> <li>26.1 異分野間の連携をはかったパイプラインの構築</li> </ol> </li> <li>27 in silico 解析の限界</li> <li>28 マイクロアレイ解析の概要 <ol style="list-style-type: none"> <li>28.1 大規模なハイブリダイゼーションによる発現解析</li> </ol> </li> <li>29 マイクロアレイ解析の理論的根拠 <ol style="list-style-type: none"> <li>29.1 セントラルドグマ</li> </ol> </li> <li>30 マイクロアレイ解析に関する生物学的な注意点</li> <li>31 代表的なマイクロアレイ技術 <ol style="list-style-type: none"> <li>32 1) オリゴタイプマイクロアレイ <ol style="list-style-type: none"> <li>32.1 Affymetrix 社の GeneChip (DNA チップ)</li> </ol> </li> <li>33 オリゴタイプマイクロアレイ作製法 <ol style="list-style-type: none"> <li>33.1 Ink-jet printing 技術</li> </ol> </li> <li>34 2) cDNA タイプマイクロアレイ <ol style="list-style-type: none"> <li>34.1 スポットタイプマイクロアレイ</li> </ol> </li> <li>35 cDNA タイプマイクロアレイ作製法 <ol style="list-style-type: none"> <li>35.1 スタンフォードタイプ</li> </ol> </li> <li>36 3) SAGE(Serial Analysis of Gene Expression)法 <ol style="list-style-type: none"> <li>36.1 mRNA プールにおける定量的な測定が可能</li> </ol> </li> </ol> </li> </ol>

37	オリゴタイプマイクロアレイと cDNA タイプマイクロアレイの比較
38	マイクロアレイ解析データの誤差(1)
38.1	マイクロアレイ作製、実験、画像処理、蛍光色素
39	スポットの認識と数値化(バックグラウンドの数値化)
40	マイクロアレイ解析データの誤差(2)
40.1	PMT(Photomultiplier)(光電子倍増管)と蛍光イメージの取り込み
41	実験デザイン
41.1	実験デザイン空間とは?
41.2	実験デザイン空間をいかに決定するか?
41.3	実験計画法の必要性
42	マイクロアレイの繰り返し実験と結果の再現性
43	そもそも繰り返し実験とは?
44	マイクロアレイ解析の典型的な目的と疑問
45	マイクロアレイデータの正規化(normalization)
45.1	正規化とは
45.2	対数変換と正規分布
45.3	なぜ、正規化が必要なのか?
45.4	線形回帰と主成分分析による正規化
45.5	MA-plot と lowess 関数による正規化
45.6	MA-plot と lowess 関数による正規化の例とまとめ
46	SMD: Stanford Microarray Database
47	SMD のデータを閲覧する(1)
48	SMD のデータを閲覧する(2)
48.1	Clickable image, View Array Image and Grids
49	SMD のデータを閲覧する(3), (4)
49.1	View Array Details, Plot Array Data
50	R - 統計解析ソフトウェア
51	R の基本的な使い方
51.1	CSV ファイルの読み込み、データ列の表示、ヘルプ
52	R による plot 図の表示
53	R による MA-plot での正規化と Lowess 関数による局所重み付き多項式回帰
54	参考文献

## (5) 第5回

日時	2004年11月17日
科目名	発現プロファイル解析2
時間・コマ数	3時間(1コマ)
講師名	荻島 創一
講義概要	<ol style="list-style-type: none"> <li>1 第1回ゲノム医療情報シンポジウム</li> <li>2 PSB 2005</li> <li>3 5<sup>th</sup> International Conference on Systems Biology</li> <li>4 講義内容</li> <li>5 フィルタリング</li> <li>6 発現変化の有意性 <ol style="list-style-type: none"> <li>6.1 パラメトリック検定- Welch の t 検定</li> <li>6.2 パラメトリック検定- 分散分析(ANOVA)</li> <li>6.3 ノンパラメトリック検定- Wilcoxon/Mann-Whitney の順位和検定</li> </ol> </li> <li>7 多重検定に対する補正</li> <li>8 主成分分析(PCA)</li> <li>9 クラスタ解析 <ol style="list-style-type: none"> <li>9.1 階層化クラスタ化法</li> <li>9.2 K-means クラスタ化法</li> <li>9.3 自己組織化マップ(SOM)</li> <li>9.4 ユークリッド距離、コサイン係数距離、ピアソン距離</li> </ol> </li> <li>10 サポートベクターマシン(SVM)によるクラス識別</li> <li>11 発現プロファイル解析のためのソフトウェア</li> <li>12 R - 統計解析ソフトウェア(再掲)</li> <li>13 Rの基本機能</li> <li>14 作業ディレクトリの設定</li> <li>15 データのファイル出力</li> <li>16 演習データ(1): Diffuse large B-cell lymphoma (DLBCL)における遺伝子発現データ</li> <li>17 Rの基本的な使い方(再掲・修正) <ol style="list-style-type: none"> <li>17.1 CSVファイルの読み込み、データ列の表示、ヘルプ</li> </ol> </li> <li>18 マイクロアレイデータの前処理 <ol style="list-style-type: none"> <li>18.1 intensityの分布のヒストグラム表示</li> <li>18.2 2チャンネル(Cy5/Cy3)間のintensityの分布の比較</li> <li>18.3 plot図の表示(前回の復習)</li> <li>18.4 両対数のplot図の表示</li> <li>18.5 最小二乗法での線形回帰</li> <li>18.6 主成分分析によるplot図の眺め</li> <li>18.7 MA-plotでの正規化とlowess関数(前回の復習)</li> </ol> </li> <li>19 演習データ(2): 4個の遺伝子と6つの症例および病期の遺伝子発現データ(演習用サンプルデータ)</li> <li>20 データの読み込みと表示</li> <li>21 有意な発現変化を示す遺伝子の同定 <ol style="list-style-type: none"> <li>21.1 Welchのt検定</li> <li>21.2 分散分析(ANOVA) <ol style="list-style-type: none"> <li>21.2.1 t検定と分散分析(ANOVA)での評価の比較</li> </ol> </li> <li>21.3 Wilcoxonの順位和検定</li> <li>21.4 主成分分析</li> <li>21.5 Bioconductor</li> <li>21.6 Bioconductorによる解析例</li> <li>21.7 必要なライブラリのロード <ol style="list-style-type: none"> <li>21.7.1 Biobase, annotate, genefilter, golubEsets, cluster, ellipse, lattice, XML, e1071のロード</li> </ol> </li> </ol> </li> <li>22 演習データ(3): ALLとAMLの遺伝子発現データ</li> <li>23 解析を始めるまえに <ol style="list-style-type: none"> <li>23.1 概要の閲覧</li> <li>23.2 発現データの表示</li> </ol> </li> </ol>

	23.3	表現型データの表示
	23.4	最初の3個の遺伝子と10個のチップのデータの抽出
	23.5	発現レベルのデータの抽出
24		フィルタリング
	24.1	発現レベルフィルタリング
	24.2	バリエーションフィルタリング
25		マイクロアレイデータの前処理
	25.1	log 変換
	25.2	標準化
26		相関係数の計算と相関係数行列の表示
27		相違関係行列のレベルプロット表示
28		有意な発現変化を示す遺伝子の同定
	28.1	ALL と AML の2群についての t 検定
	28.2	同定された遺伝子に関する各統計量
	28.3	相関係数と計算と相関係数行列の表示
29		多次元尺度法による表示
	29.1	2次元 (k=2) の場合
	29.2	3次元 (k=3) の場合
30		クラスタ解析
	30.1	階層的クラスタ化法(average linkage の場合)
	30.2	階層的クラスタ化法(single linkage の場合)
	30.3	階層的クラスタ化法(complete linkage の場合)
	30.4	PAM クラスタ化法(2つのクラスタに分けた場合)
		PAM クラスタ化法(3つのクラスタに分けた場合)
	30.6	K-means クラスタ化法(2つのクラスタに分けた場合)
	30.7	K-means クラスタ化法(3つのクラスタに分けた場合)
	30.8	何個の K-means クラスタでクラスタ化すればよいのか?
	30.9	クラスタ解析で用いられるそのほかのソフトウェア
	30.10	Cluster ソフトウェアを用いたクラスタ解析
	30.11	TreeView によるクラスタ解析結果の表示
31		SVM によるクラス識別
	31.1	識別のための超平面の発見
	31.2	トレーニングデータでのクラス識別のテスト
	31.3	テストデータでのクラス識別のテスト
32		PubMed データベースと連携する
33		参考文献

## (6) 第6回

日時	2004年11月24日
科目名	発現プロファイル解析3
時間・コマ数	3時間(1コマ)
講師名	荻島 創一
	<ol style="list-style-type: none"> <li>1 講義内容</li> <li>2 演習データ(3): ALLとAMLの遺伝子発現データ</li> <li>3 必要なライブラリのロード <ol style="list-style-type: none"> <li>3.1 Biobase, annotate, genefilter, golubEssets, cluster, ellipse, lattice, XML, e1071のロード多重検定に対する補正</li> </ol> </li> <li>4 解析するデータの下準備(再掲・修正) <ol style="list-style-type: none"> <li>4.1 データのロード、フィルタリング、log変換、標準化</li> </ol> </li> <li>5 相関係数での眺め(前回未終了分) <ol style="list-style-type: none"> <li>5.1 相関係数の計算と相関係数行列の表示</li> <li>5.2 相関係数行列のレベルプロット表示</li> </ol> </li> <li>6 有意な発現変化を示す遺伝子の同定 <ol style="list-style-type: none"> <li>6.1 ALLとAMLの2群についてのt検定</li> <li>6.2 同定された遺伝子に関する各統計量</li> <li>6.3 相関係数の計算と相関係数行列の表示</li> </ol> </li> <li>7 多次元尺度法による表示 <ol style="list-style-type: none"> <li>7.1 2次元(k=2)の場合</li> <li>7.2 3次元(k=3)の場合</li> </ol> </li> <li>8 クラスタ解析 <ol style="list-style-type: none"> <li>8.1 階層的クラスタ化法(average linkageの場合)</li> <li>8.2 階層的クラスタ化法(single linkageの場合)</li> <li>8.3 階層的クラスタ化法(complete linkageの場合)</li> <li>8.4 PAMクラスタ化法(2つのクラスタに分けた場合)</li> <li>8.5 PAMクラスタ化法(3つのクラスタに分けた場合)</li> <li>8.6 K-meansクラスタ化法(2つのクラスタに分けた場合)</li> <li>8.7 K-meansクラスタ化法(3つのクラスタに分けた場合)</li> <li>8.8 何個のクラスタでクラスタ化すればよいか?</li> <li>8.9 クラスタ解析で用いられるそのほかのソフトウェア</li> <li>8.10 Clusterソフトウェアを用いたクラスタ解析</li> <li>8.11 TreeViewによるクラスタ解析結果の表示</li> </ol> </li> <li>9 SVMによるクラス識別(前回未終了分) <ol style="list-style-type: none"> <li>9.1 識別のための超平面の発見</li> <li>9.2 トレーニングデータでのクラス識別のテスト</li> <li>9.3 テストデータでのクラス識別のテスト</li> </ol> </li> <li>10 PubMedデータベースと連携する(前回未終了分)</li> <li>11 オントロジー、データベース連携、ネットワーク推定、シミュレーションへ</li> <li>12 -omeデータの統合と疾患情報との連携(再掲・修正)</li> <li>13 Gene Ontology(再掲)</li> <li>14 Gene Ontologyの例 <ol style="list-style-type: none"> <li>14.1 SGD(Saccharomyces Genome Database)</li> </ol> </li> <li>15 O B F - Open Bioinformatics Foundation(再掲)</li> <li>16 DAS(Distributed Annotation Server)</li> <li>17 文献情報の自然言語処理によるデータベース <ol style="list-style-type: none"> <li>17.1 自動構築(再掲)</li> </ol> </li> <li>18 Biocompass <ol style="list-style-type: none"> <li>18.1 NEC社による網羅的な検索が可能な文献検索ツール</li> </ol> </li> <li>19 医学用語のオントロジー</li> <li>20 医学用語と生物学用語のオントロジーの融合の試み(再掲)</li> <li>21 システムバイオロジーの基本戦略(再掲)</li> <li>22 代表的なパスウェイデータベース</li> <li>23 KEGG <ol style="list-style-type: none"> <li>23.1 Pathway Mapsをブラウズする</li> </ol> </li> </ol>

	23.2 Ortholog Tables をブラウズする
	23.3 酵素/化合物/遺伝子で検索する
	23.4 酵素/化合物/遺伝子を色付けして表示する
	23.5 経路探索する
24	遺伝子発現データからの遺伝子調節
	24.1 ネットワークの推定 (前々回未説明)
25	代表的な細胞のシミュレーション (前々回未説明)
26	E-Cell (前々回未説明)
27	北野共生システムプロジェクト
28	SBML (前々回未説明)
29	シミュレーションをする
	29.1 SBW/Jarnac/JDesigner
30	SBW/Jarnac/JDesigner でシミュレーション
	30.1 インストールとセットアップ
	30.2 モデルを Load する
	30.3 時系列シミュレーション
	30.4 ノード (物質)、エッジ (反応) の情報をブラウズする
	30.5 安定性を解析する
	30.6 条件を変えてシミュレーションするほか
	30.7 モデルをデザインする
31	ショウジョウバエの概日リズムのシステム解析
32	京大細胞・生体シミュレータプロジェクト (前々回未説明)
	32.1 FDA 治験に耐えうるシミュレータの開発へ
	32.2 膵細胞シミュレータ
33	システム生物学の医療・創薬への応用 (前々回未説明)
34	シグナル伝達系の乱れとしての疾患
35	シグナル伝達系のモデリング
36	血管構築のモデリングとシミュレーション
37	Physiome Project
38	Entelos
	38.1 in silico R&D による効率的な創薬の試み
39	The Virtual Heart Project
40	Pharmacogenetics/genomics との連携 (前々回未説明)
	40.1 パーソナライズド・メディスンへ
41	参考文献

## (7) 第7回

日時	2004年12月4日
科目名	多型解析(SNP)
時間・コマ数	3時間(1コマ)
講師名	今西 規
講義概要	<p>1 講義内容</p> <p>2 実習内容</p> <p>3 1. ゲノムの変異(SNP・ハプロタイプとは)</p> <p>4 遺伝的多型とは</p> <p>5 Figure2.7 Schematic representation of the dynamics of gene substitution for (a) advantageous and (b) neutral mutations</p> <p>6 ゲノム多型の種類</p> <p>7 表3 SNPの表現型への影響</p> <p>8 図2 SNPの分類と意義</p> <p>9 Figure1.12 Types of substitution mutations in a coding region</p> <p>10 2. 突然変異の起こり方</p> <p>11 偽遺伝子における塩基置換パターン</p> <p>12 塩基置換パターンの特徴</p> <p>13 SNPの祖先型対立遺伝子</p> <p>14 SNPの塩基置換パターン</p> <p>15 OMIM(遺伝性疾患に関するあらゆる臨床情報を提供してくれるデータベース)に登録されているORFアミノ酸置換情報の分析</p> <p>16 OMIM allelic variant Example</p> <p>17 Amino acid substitution pattern of OMIM Allelic variant</p> <p>18 H-Invitational プロジェクト</p> <p>19 H-インビテーションショナル・データベース</p> <p>20 H-インビテーションショナル論文における主な発見</p> <p>21 3. dbSNPをはじめとする多型データベース</p> <p>22 SNP関連データベース</p> <p>23 ハプロタイプ関連データベース</p> <p>24 NCBIのdbSNP</p> <p>25 dbSNPの入力情報</p> <p>26 NCBIのdbSNPのデータ</p> <p>27 dbSNP build 115 (June 2003) statistics</p> <p>28 NCBI dbSNP Build 121(July 2004)</p> <p>29 4. ヒトゲノムの多様性解析</p> <p>30 SNPの分布(ヒト第2染色体)</p> <p>31 マイクロサテライトDNAの分布(ヒト第2染色体)</p> <p>32 染色体領域の進化の歴史と変異の量</p> <p>33 SNP, マイクロサテライトDNAと塩基組成(G+C含量)の相関染色体バンドと多様性の関係</p> <p>34 染色体バンドと多様性の関係</p> <p>35 5. 疾患との関係</p> <p>36 さまざまな遺伝性疾患の原因</p> <p>37 例1: 鎌状赤血球貧血</p> <p>38 図4.21 グロビン遺伝子の単一ヌクレオチド置換えによるグロビンの単一アミノ酸置換</p> <p>39 超優性淘汰によるHbSの維持</p> <p>40 NCBIのOMIM(遺伝性疾患情報)</p> <p>41 実習課題</p> <p>42 課題1. 公共データベースからのSNP情報の検索</p> <p>43 課題1. 公共データベースからのSNP情報の検索</p> <p>44 課題1. 公共データベースからのSNP情報の検索</p> <p>45 課題2. SNPハプロタイプの推定</p> <p>46 集団データからSNPハプロタイプを推定する</p> <p>47 遺伝統計学の初歩</p> <p>48 カイ二乗検定</p>

49	ハーディー・ワインバーグ平衡
50	ハーディー・ワインバーグ平衡の検定
51	カイ二乗検定による独立性の検定
52	Fisher の正確確率検定
53	Bonferroni の補正
54	実習課題
55	3 ( 1 ) ハーディー・ワインバーグ平衡の検定を行う
56	3 ( 2 ) 疾患感受性 S N P の発見のための検定を行う
57	実習課題の解答
58	解答 2
59	解答 3 ( 1 )
60	解答 3 ( 2 )

## ( 8 ) 第 8 回

日時	2004 年 12 月 15 日
科目名	プロテオーム解析とタンパク質相互間解析
時間・コマ数	3 時間 ( 1 コマ )
講師名	櫻井 仁美
	1 生物の解析 2 プロテオームとプロテオミクス 3 プロテオミクスの研究戦略 4 蛋白質の発現解析 5 蛋白質の分離と同定方法 6 2D-PAGE の原理と実際-1 7 2D-PAGE の原理と実際-2 8 2D-PAGE の原理と実際-3 9 スポット同定のための質量分析法 10 MALDI-TOF/MS の原理 11 質量分析による解析の実際 12 PMF による蛋白質の同定 13 PMF による蛋白質の同定 ~ 実習 ~ 14 Mascot-1 15 Mascot-2 16 Mascot-3 17 Mascot-4 18 Mascot-5 19 例：大腸菌から精製した蛋白質 20 MASCOT 解析結果の見方 1 21 MASCOT 解析結果の見方 2 22 MASCOT 解析結果の見方 3 23 例：大腸菌から精製した蛋白質 ( 応用 ) 23.1 応用-1 23.2 応用-2 23.3 応用-3 24 蛋白質データベース 25 蛋白質データベース ~ 実習 ~ 25.1 ExPASy 26 2D-PAGE データベース-1 27 2D-PAGE データベース-2 28 条件で変動する蛋白の同定 29 蛋白質の相互作用解析 30 蛋白質の相互作用解析法 31 酵母 Two-Hybrid 法の原理 32 プルダウン法の原理-1 33 プルダウン法の原理-2 34 プルダウン法の原理-3 35 タグ一覧 36 蛋白質相互作用データベース ~ 実習 ~ 36.1 蛋白質相互作用データベース -1 36.2 蛋白質相互作用データベース -2 37 蛋白質相互作用データベース ~ 実習 : BIND ~ 37.1 BIND-1 37.2 BIND-2 38 蛋白質相互作用データベース ~ 実習 : PathCalling ~ 38.1 PathCalling-1 38.2 PathCalling-2 38.3 PathCalling-3 39 本日のまとめ 40 プロテオミクスの研究戦略

### C. アノテーター育成実証講座 終了テスト

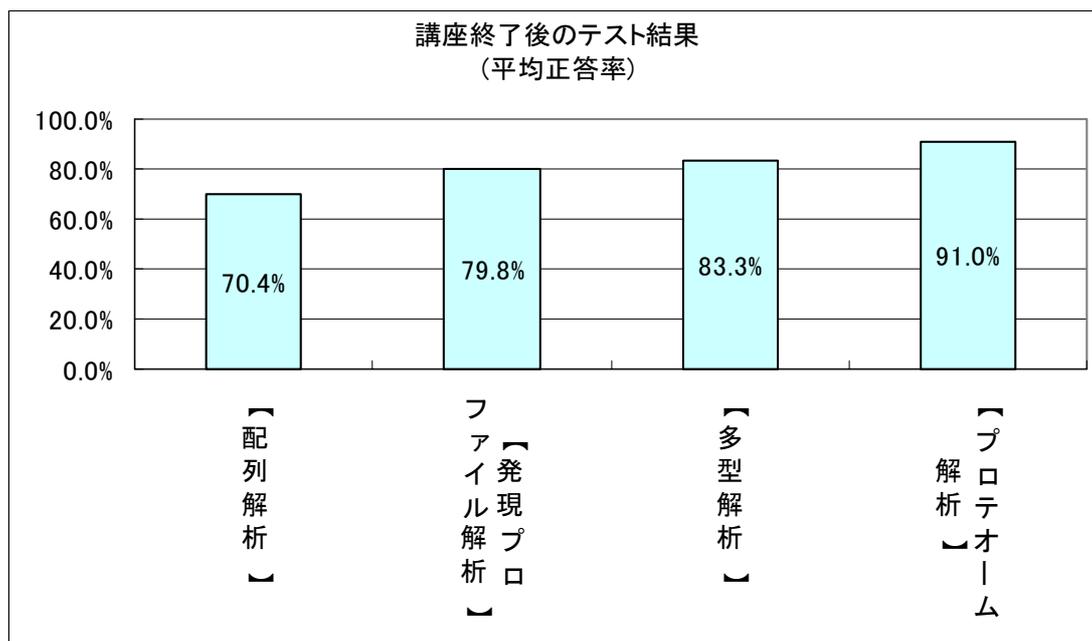
全8回の講座が終了した後、講義の理解度を見るための復習テストを実施した。

問題は4つの選択肢の中から選ぶ選択式で、8回の講座の一部を欠席した場合には受講分野についてのみ解答することとした。問題および解答用紙は、電子メールで受け渡しを行なった。

全講座を受講し、全50問をすべて回答した受講者は14人中11名であった。この11名の正答率は78.0%(平均正答数39問)、最高90.0%、最低60.0%であった。

分野	出題数	回答者	平均点	平均正答率	最高	最低
【配列解析】	20	12	14.1	70.4%	95.0%	50.0%
【発現プロファイル解析】	18	11	14.4	79.8%	94.4%	50.0%
【多型解析】	6	13	5.0	83.3%	100.0%	66.7%
【プロテオーム解析】	6	13	5.5	91.0%	100.0%	66.7%
全体	50	11	39	78.0	90.0%	60.0%

分野ごとの正答率グラフ



全講座が終了した後に出题したため、【多型解析】、【プロテオーム解析分野】など後半の学習分野からの出題は正答率がやや高く、受講から約2ヶ月経過している【配列解析】分野の正答率はやや低いという結果になった。いずれにせよ、どの分野も70%を超える高い正答率となった。

次のページより、出題された50問、及び解答一覧を掲載する。

【配列解析】

問1 ある塩基配列を得た時に様々な検索・解析を実行し、その配列の機能推定を行うことがある。次のデータベースのうち、塩基配列を query (問い合わせ) にして検索が実行できないデータベースはどれか。

- 1 . MGI
- 2 . FlyBase
- 3 . Ensembl
- 4 . SOSUI

問2 配列相同性検索プログラム BLAST を実行する際に、query (問い合わせ) と subject (参照) の配列のタイプによって使用するプログラムを選択するが次のうち、正しい組み合わせはどれか。

プログラム	query (問い合わせ)	subject (参照)
1 . TBLASTN	アミノ酸配列	塩基配列を翻訳したアミノ酸配列
2 . BLASTN	塩基配列	アミノ酸配列
3 . BLASTP	塩基配列	アミノ酸配列
4 . TBLASTX	アミノ酸配列	塩基配列を翻訳したアミノ酸配列

問3 ゲノムは生物の遺伝情報を全て含んだ貯蔵庫である。近年の配列読み取り技術の進歩に伴い、多数の生物のゲノム配列が明らかされ、またさらにゲノム解析が進行中の生物も数多く存在する。1995 年に自立生物として最初のゲノム配列が決定された。その生物は次のうちどれか。Escherichia coli K12 MG1655

- 2 . Saccharomyces cerevisiae
- 3 . Mus musculus
- 4 . Haemophilus influenzae

問4 以下のモデル生物とデータベースの組み合わせのうち、間違っているのはどれか。

- 1 . *Mus musculus* MGI
- 2 . *Caenorhabditis elegans* Wormbase
- 3 . *Drosophila melanogaster* SGD
- 4 . *Oryza sativa japonica* RGP

問5 データベース検索を行う際に、検索したい項目や内容の関係を組み合わせることにより、効率的に検索を実行できる。下記の論理演算子の内容として、妥当ものはどれか。

- 1 . 論理積 OR
- 2 . 論理和 XOR
- 3 . 論理否定 NOT
- 4 . 排他的論理和 AND

問6 ゲノム配列や cDNA (complementaryDNA) といった塩基配列上にタンパク質コード領域が存在するかを調べる際に注意すべき事柄として、妥当なものはどれか。

- 1 . 予測プログラムで予測プログラムを用いて、タンパク質コード領域を調べる時には複数のプログラムを用いる必要はなく、一つのプログラムだけを用いればよい。
- 2 . タンパク質コード領域を調べるにあたっては、予測プログラムの結果だけではなく、既存の遺伝子や転写物上のアミノ酸配列との配列相同性も調べたほうが良い。

3. 既存のアミノ酸配列と配列相同性が全く認められない転写物は実験上、誤った取れた転写物であるので、今後の解析から排除すべきである。
4. 得られたタンパク質コード領域と既存のタンパク質データベースとの配列相同性検索によって得られた結果が同じである時は、その領域が間違いなくタンパク質コード領域であるので、実験的な確認を行う必要は全くない。

問7 文献データベースに関する記述のうち、最も妥当なものはどれか。

1. MEDLINE は NCBI が作成している医療分野用文献データベースである。
2. PubMed は NCBI で提供されている統合検索サービスの対象に含まれてない。
3. PubMed には MEDLINE に登録されている文献が含まれている。
4. MeSH とは Medical Sequence Heading の略称であり、PubMed に登録されている雑誌のインデキシングに用いられる語彙のことである。

問8 ある真正細菌のゲノム配列(約 3M bp)の完全長を決定した。遺伝子予測プログラムを用いて ORF の推定を行おうとする場合、最も適しているプログラムを一つ選べ。

1. Glimmer
2. GlimmerHMM
3. GENESCAN
4. FGENES

問9 遺伝子予測プログラム Glimmer を用いる際に予測の正確性を高めるため、最も適していると考えられる手法を一つ選べ。

1. タンパク質コード領域でない rRNA, tRNA, non-coding RNA 領域を予測し、それらの領域をマスキングしてから Glimmer を実行した。
2. rRNA, tRNA, non-coding RNA 領域に加え、相同性検索ですぐに分かるようなタンパク質コード領域は予測する必要がないのでそれらの領域をマスキングしてから Glimmer を実行した。
3. rRNA, tRNA, non-coding RNA 領域に加え、外来性因子(トランスポゾンやファージ由来と考えられるゲノム領域)の領域を事前にマスキングして Glimmer を実行した。
4. Glimmer のアルゴリズム上、長い ORF が存在すると正確度が落ちるため、rRNA, tRNA, non-coding RNA 領域に加え、3,000bp を超えるオープンリーディングフレームをマスキングし Glimmer を実行した。

問10 ある細菌からショットガンシーケンス法により 10,000bp の DNA 断片の塩基配列を決定した。塩基配列をクエリー(問い合わせ)配列に用いて、BLAST 検索でこの DNA 断片内にある遺伝子を見つけようとする場合、最も適したプログラムを一つ選べ。

1. blastx
2. blastp
3. tblastn
4. blastx と blastn を両方行う

問11 アミノ酸配列長が 85 残基を超える場合、NCBI(National Center for Biotechnology Information)の BLASTP では、どのマトリックスの利用を推奨しているか。

- 1 . BLOSUM62
- 2 . PAM250
- 3 . PAM70
- 4 . BLOSUM80

問12 BLASTP の出力結果を評価する場合に最も適した内容を選べ。

1. E-value が同じであれば Identities も同じであるため、すべて E-value に基づいて複数の相同性を判定すればよい。
2. 相同性の内容は E-value だけでなく、Alignment の内容も考慮しなければならない。
3. Identities で示される%値が高いほど相同性が高いので、この値を検索することで有意な相同性を示す結果のみを素早く検索することができる。
4. Identities が 100%なら E-value は 0 であるため、E-value=0 を対象に結果を検索すれば完全一致したタンパク質のみを拾い上げることができる。

問13 ある 2 株の細菌間では、染色体の塩基配列が非常に類似していることが分かっている。これら 2 株の塩基配列どうしを blast プログラムで高速比較する場合の手段として最適なものを一つ選べ。

1. rpsblast を行う
2. blastx を行う
3. tblastx を行う
4. megablast を行う

問14 a, t, g, c のみで構成される(曖昧な配列を含まない)塩基配列をクエリーに用いて Swiss-Prot データベースを対象に BLASTX を行ったところ、出力されるアラインメントファイル中に X が含まれていた。これについて正しい説明はどれか。

1. Low complexity フィルターオプションが ON になっていたため、タンパク質データベース側(Swiss-Prot 側)の解析対象とならない領域が X に置換されて表示された。
2. Low complexity フィルターオプションが ON になっていたため、クエリー配列側の解析対象とならない領域が X に置換されて表示された。
3. Low complexity フィルターオプションが ON になっていたため、解析対象とならないタンパク質データベース側(Swiss-Prot 側)とクエリー配列の両方の領域が X に置換されて表示された。
4. クエリーの塩基配列からアミノ酸への翻訳を行う際、特定のアミノ酸への訳が不可能であったため、クエリー配列側の対応する配列が X で置換された。

問15 相同遺伝子に関する次の説明で誤った説明をしているものを選べ。

- 1 . 遺伝子重複により分かれた遺伝子ペアをパラログと呼ぶ。
- 2 . 種分岐により分かれた遺伝子をオルソログと呼ぶ。
- 3 . 生物種間で相同な遺伝子ペアをオルソログと呼ぶ。
- 4 . パラログな遺伝子ペアも、オルソログな遺伝子ペアもホモログな遺伝子ペアと呼べる。

問16 系統樹作成法でないものはどれか？

- 1 . Neighbor-Joining
- 2 . Maximum Parsimony
- 3 . Maximum Likelihood
- 4 . Mann-Whitney

問17 次の学名と種名の対応で間違っているものは何か？

学名	生物種名
1 . <i>Saccharomyces cerevisiae</i>	出芽酵母
2 . <i>Arabidopsis thaliana</i>	シロイヌナズナ
3 . <i>Dugesia japonica</i>	ライス
4 . <i>Ciona intestinalis</i>	ホヤ

問18 NCBI に存在する各種データベース、及びプログラムの説明で正しいものはどれか？

- 1 . Tax Browser は学名より生物種を探るために利用される。
- 2 . ORF Finder はアミノ酸配列から DNA 配列を推定するソフトウェアの一つである。
- 3 . Genome Information Broker はゲノム配列を収めたデータベースである。
- 4 . Entrez は文献検索エンジンである。

問19 CLUSTALX に関する以下の記述の内、正しいものはどれか？

- 1 . .phr という拡張子のファイルでないと配列を読み込むことが出来ない。
- 2 . UPGMA による系統樹作成のみが行える。
- 3 . 同一名の配列を読み込むことは出来ない。
- 4 . アラインメント作業を実行することで、ギャップのないアラインメントを作成することが出来る。

問20 配列解析に関する以下の記述で最も正しいものを選び。

- 1 . 系統樹を作成し、遺伝子重複、もしくは種分岐による遺伝子分岐の推定を行った。
- 2 . 2 生物種間の 2 つの遺伝子配列が相同なのでオルソログと判断した。
- 3 . 遺伝子配列の名前から配列の相同性の推定を行った。
- 4 . 配列の相同性はなかったが、系統樹を作成し、遺伝子の進化過程の推定を行った。

#### 【発現プロファイル解析】

問21 マイクロアレイ解析の概要について述べたものとしてもっともふさわしくないものを以下の 4 つの選択肢のなかから選べ。

- 1 . 従来からハイブリダイゼーションを基本原理とする解析はなく、ゲノム情報の急増とハイスループット技術の要請がマイクロアレイ解析を生んだ
- 2 . 網羅的な解析により研究対象に関する事前の知識のバイアスに依存せずに、新しい仮説を導き出せるようになった
- 3 . 症例について、新たに遺伝子発現の有意差のあるサブグループに分けて診断し、適切な治療・投薬・予後予測が可能になった
- 4 . これまでの典型的な臨床研究では数千から数万の症例から数百の変数が解析されたが、マイクロアレイ研究では数十から数百の症例から数千の変数が解析される

問22 マイクロアレイ解析の特徴を述べたキーワードとしてもっともふさわしくないものを以下の4つの選択肢のなかから選べ。

1. 並列性
2. 網羅性
3. 大規模性
4. 高再現性

問23 マイクロアレイ解析におけるパイプライン処理について述べたものとしてもっともふさわしくないものを以下の4つの選択肢のなかから選べ。

1. 通常のパイプライン処理は、大きく分けて、適切なサンプルの採取、RNAの抽出、チップへのハイブリダイゼーション、チップの画像処理、クラスタ解析、データマイニング、有意性の評価、仮説の導出、生物学的な検証から構成される
2. RNAの抽出、チップへのハイブリダイゼーション、チップの画像処理は実験手技上のミスの影響を受けやすい
3. バイオインフォマティクスが関わる部分はクラスタ解析、データマイニング、有意性の評価のみであり、パイプラインの最初のデザインから関与する必要はない
4. データマイニングにより得られる仮説は実験により検証しなければならないことがほとんどである

問24 マイクロアレイ解析に関する生物学的な注意点としてもっともふさわしくないものを以下の4つの選択肢のなかから選べ。

1. mRNAのセットであるトランスクリプトームはmRNAのプールとして考えられる
2. mRNAは数百塩基/分の速さで転写されうるが、それよりも遅い場合もある
3. mRNAの濃度はそれがコードするタンパク質の濃度と正確に対応している
4. 遺伝子多型によりある1つの遺伝子型のmRNAのみが検出されることがある

問25 代表的なマイクロアレイ技術としてもっともふさわしくないものを以下の4つの選択肢のなかから選べ。

1. cDNAタイプマイクロアレイ(スポットタイプマイクロアレイ)
2. オリゴタイプマイクロアレイ(Affymetrix社のGeneChip)
3. SAGE(Serial Analysis of Gene Expression)
4. SAGE(Sequence Analysis of Gene Expression)

問26 オリゴタイプマイクロアレイの特徴としてもっともふさわしくないものを以下の4つの選択肢のなかから選べ。

1. プローブの配列は合成するため自由に設計することができる
2. プローブに数千merの長い配列を用いることができる
3. コンピュータチップの製造技術である光リソグラフィ処理によりチップの高密度化を実現している
4. 完全マッチからミスマッチを差し引くことでその遺伝子発現を検出する

問27 cDNAタイプマイクロアレイデータのノイズの原因について述べたものとしてもっともふさわしくないものを以下の4つの選択肢のなかから選べ。

1. サンプルの調製におけるmRNAの調製、転写(酵素など)、標識(種類、プロトコール)、増幅(PCRのプロトコール)
2. チップの作製におけるスポットティングの位置、量、固定のばらつき

3. ハイブリダイゼーションにおける条件、チップの不均一性
4. 画像処理における特異的ハイブリダイゼーションのゲイン設定(PMT)、認識スポットのずれ

問28 マイクロアレイデータの前処理で対数変換が必要な理由としてもっともふさわしいものを以下の4つの選択肢のなかから選べ。

1. 歴史的な経緯により対数変換をすることが必要となった
2. 蛍光強度の比率の分布は対数正規分布で表現できる場合が多いため
3. 対数変換により発現変化の有意性を検定することができるため
4. 対数変換により発現パターンの類似性に基づくクラスタ解析をすることができるため

問29 マイクロアレイデータの正規化に、MA プロットと lowess 関数による正規化があるが、MA プロットについて述べたものとしてもっともふさわしいものを以下の4つの選択肢のなかから選べ。

1. MA プロットの MA とは MicroArray の略である
2. 発現レベルの和に対して、発現比の log をプロットしたものである
3. 発現レベルの差に対して、発現比の log をプロットしたものである
4. 発現レベルの和に対して、発現差の log をプロットしたものである

問30 オープンソースの統計解析ソフトウェアに R があるが、これは統計解析で用いられているある言語を移植したものとみなせる。そのある言語として正しいものを以下の4つの選択肢のなかから選べ。

1. S 言語
2. T 言語
3. U 言語
4. V 言語

問31 統計解析ソフトウェア R において、表データ x の CH1MEAN と CH2MEAN の2つの列についてプロット図を表示する書式として正しいものを以下の4つの選択肢のなかから選べ。

1. `plot(x$CH1MEAN, x$CH2MEAN)`
2. `plot(x->CH1MEAN, x->CH2MEAN)`
3. `plot{x->CH1MEAN, x->CH2MEAN}`
4. `plot(x$CH1MEAN, x$CH2MEAN)`

問32 7000個の遺伝子で偽陽性を1つ含む確率が25%であるように Bonferroni の多重検定に対する補正を行った場合、P 値の棄却条件としてもっともふさわしいものを以下の4つの選択肢のなかから選べ。

1.  $3 \times 10^{-3}$
2.  $3 \times 10^{-4}$
3.  $3 \times 10^{-5}$
4.  $3 \times 10^{-6}$

問33 階層型クラスタ化法と K-means クラスタ化法について述べたものとしてもっともふさわしくないものを以下の4つの選択肢のなかから選べ。

1. 階層化クラスタ化法では、すべての遺伝子間の距離を計算し、互いにもっとも近い距離を示す遺伝子をグループとし、階層的にグループ分けし、クラスタ化する
2. 階層化クラスタ化法において、各クラスタ間の距離をもっとも近い点で定義する

ことを Average Linkage Clustering とよぶ

- 3 . K-means クラスタ化法では、すべての遺伝子間の距離を計算し、いくつかのクラスターをあらかじめ用意し、各クラスターの中心との距離がもっとも近いクラスターに割り当てることでクラスタ化する
- 4 . K-means クラスタ化法において、各クラスターの中心をセントロイドとよぶ

問34 SVM によるクラス識別について述べたものとしてもっともふさわしいものを以下の4つの選択肢のなかから選べ。

- 1 . SVM はサポートベストマッチマシンの略であり、機械学習の手法のひとつである
- 2 . 機械学習には教師あり学習と教師なし学習の2種類があるが、SVM は教師なし学習である
- 3 . SVM は教師なしニューラルネットワーク、自己組織化マッピングと非常によく似ている
- 4 . SVM は高次元の特徴空間においてクラスを線形判別するための超平面を発見する

問35 オントロジーについて述べたものとしてもっともふさわしくないものを以下の4つの選択肢のなかから選べ。

- 1 . 遺伝子のオントロジーには Gene Ontology があり、これは OBO (Open Bioinformatics Ontologies) による機能アノテーションのための階層的な機能分類である
- 2 . Gene Ontology には Molecular function、biological process、cellular component の3つの大分類がある
- 3 . OBO はゲノムが解読された各生物種の遺伝子への機能アノテーションを行っている
- 4 . 医学用語のオントロジーの整備は完了しており、代表的なものに UMLS Biomedical Ontology がある

問36 発現プロファイルからの遺伝子調節ネットワークの推定について述べたものとしてもっともふさわしくないものを以下の4つの選択肢のなかから選べ。

- 1 . 遺伝子調節ネットワークの表現方法として、ブーリアンネットワーク、またはベイジアンネットワークが用いられることが多い
- 2 . ブーリアンネットワークは遺伝子の発現状態を 1 または 0 の 2 値で表現した遺伝子調節ネットワークである
- 3 . ベイジアンネットワークは遺伝子発現の状態を確率ネットワークとして表現した遺伝子調節ネットワークである
- 4 . 発現プロファイルから遺伝子調節ネットワークを必ず一意に推定することができる

問37 細胞シミュレーションについて述べたものとしてもっともふさわしくないものを以下の4つの選択肢のなかから選べ。

- 1 . 細胞シミュレーションには大別して、連立偏微分方程式を数値的に解析する連続系と離散確率モデルの2種類がある
- 2 . E-cell は連続系の細胞シミュレーションのソフトウェアであり、慶應大学の富田グループによって開発が進められている
- 3 . こうした連続系の細胞シミュレーションでは各種パラメータを同定する必要があるが、それは実験なしに簡単に行うことができる

4. 細胞シミュレーションにはさまざまなソフトウェアがあるため、共通のモデル表現系として SBML が提案され、標準化が進んでいる

問38 細胞シミュレーションソフトウェアである JDesigner/Jarnac/SBW について述べたものとしてもっともふさわしくないものを以下の4つの選択肢のなかから選べ。

1. SBW は北野共生システムプロジェクトによりシステムバイオロジーの研究の基盤として開発されたソフトウェアプラットフォームである
2. Jarnac は細胞シミュレーション time course simulation などを行うための計算エンジンである
3. JDesigner はシミュレーションしたいモデルをデザインするソフトウェアである
4. JDesigner 単体で time course simulation などの細胞シミュレーションを行うことができる

【多型解析】

問39 NCBI で作成している dbSNP というデータベースに含まれている情報を、次の中から選べ。

1. ヒトゲノムにある DNA 多型のデータ
2. ヒトのミトコンドリア DNA の制限酵素地図
3. 全生物種に由来する DNA 配列データ
4. 医学・生物学分野の文献データ

問40 タンパク質のアミノ酸配列を変化させるような DNA 多型は何か。

1. 同義置換
2. 非同義置換
3. トランジション
4. トランスバージョン

問41 ハーディー・ワインバーク平衡が成り立っている場合、2つの対立遺伝子 A と a の頻度をそれぞれ p と q ( $= 1 - p$ ) とすると、Aa というヘテロ接合の遺伝子型の頻度は次のどれになるか。

1.  $p \times p \times q \times q$
2.  $p \times q$
3.  $2 \times p \times q$
4.  $4 \times p \times q$

問42 2つの SNP サイトで構成されるハプロタイプは、理論的に何種類あるか。ただし、2つの SNP サイトはそれぞれ2種類の対立遺伝子があるものとする。

1. 2種類
2. 3種類
3. 4種類
4. 8種類

問43 2座位の連鎖不平衡の値 (D) を計算する式は次のどれか。ただし、第1の遺伝子

座の対立遺伝子頻度を  $p$  と  $1 - p$ 、第 2 の遺伝子座の対立遺伝子頻度を  $q$  と  $1 - q$ 、頻度が  $p$  と  $q$  で表される対立遺伝子の組み合わせであるハプロタイプの頻度を  $f$  とする。

- 1 .  $D = f - p q$
- 2 .  $D = f - 2 p q$
- 3 .  $D = f - p ( 1 - p ) q ( 1 - q )$
- 4 .  $D = f - p \times p \times q \times q$

問44 SNPのデータベースについて、誤った記述を選べ。

- 1 . JSNP は日本人で見つかった約 2 0 万件の SNP の情報を、インターネットで公開している。
- 2 . HGVbase はスウェーデンのカロリンスカ研究所を中心に構築されているデータベースで、約 2 8 0 万件のデータがある。
- 3 . dbSNP は NCBI で構築された世界最大の SNP データベースで、利用にはユーザー登録が必要である。
- 4 . The SNP Consortium (TSC)は、独自に発見した約 1 8 0 万件の SNP データを Cold Spring Harbor Laboratory から公開している。

#### 【プロテオーム解析】

問45 あるゲノムにコードされた蛋白質全体の呼び名はどれか？

- 1 . トランスクリプトーム
- 2 . イントラクトーム
- 3 . プロテオーム
- 4 . メタボローム

問46 一般に 2 次元電気泳動では蛋白質のどのような性質を利用して分離しているか？

- 1 . 等電点と親水性
- 2 . 分子量と等電点
- 3 . 修飾基と親水性
- 4 . 分子量と修飾基

問47 質量分析法と関係がないものはどれか？

- 1 . MALDI-TOF
- 2 . TLC
- 3 . PMF
- 4 . MASCOT

問48 蛋白質相互作用データベースでないものはどれか？

- 1 . SWISS-2DPAGE
- 2 . BIND
- 3 . MINT
- 4 . PathCalling

問49 質量分析でペプチドマスフィンガープリントに用いるデータは次の内どれか？

1. 蛋白質を一アミノ酸ずつ切断し、HPLC で分離した溶出ピーク値のセット。
2. 蛋白質を蛋白分解酵素で切断したペプチド断片の質量のセット。
3. 蛋白質を蛋白分解酵素で切断したあとのペプチド断片の 2D-PAGE の位置情報。
4. 蛋白質を蛋白分解酵素で切断したペプチド断片の等電点のセット。

問50 蛋白質の同定に関する以下の記述の内、正しいものはどれか？

1. 質量分析機では de novo アミノ酸配列決定はできない。
2. 異なるゲルの 2 次元電気泳動のパターンを比較することはできない。
3. 目的の蛋白質の配列が登録されていないとペプチドマスフィンガープリントでは蛋白質を正しく同定できない。
4. 蛋白質は翻訳後修飾が起こっても質量は変化しない。

全受講者の解答集計表

回答者/問		A	B	C	D	E	F	G	H	I	J	K	L	M	N	○	×	正答率 (%)	
【配列解析】	問1							x						x	x	9	3	75	
	問2			x										x		11	2	85	
	問3	x		x				x	x						x	7	5	58	
	問4								x	x						10	2	83	
	問5							x						x	x	10	3	77	
	問6															12	0	100	
	問7				x					x						10	2	83	
	問8	x		x					x	x						8	4	67	
	問9		x	x				x		x		x		x	x	5	7	42	
	問10		x	x	x	x	x	x	x			x		x		4	9	31	
	問11									x	x				x	x	9	4	69
	問12							x		x						x	9	3	75
	問13																13	0	100
	問14			x				x		x	x				x	x	6	6	50
	問15	x		x							x						9	3	75
	問16																12	0	100
	問17			x							x					x	9	3	75
	問18	x	x					x	x	x		x				x	5	7	42
	問19		x		x				x	x	x				x		6	6	50
	問20	x		x						x							9	3	75
【発現プロフィール解析】	問21								x							10	1	91	
	問22										x			x	x	8	3	73	
	問23															11	0	100	
	問24															11	0	100	
	問25		x											x		9	2	82	
	問26															11	0	100	
	問27						x		x		x				x	7	4	64	
	問28			x							x			x	x	8	4	67	
	問29	x							x							x	8	3	73
	問30																12	0	100
	問31														x		11	1	92
	問32					x				x		x			x		8	4	67
	問33	x	x	x	x			x		x		x			x	x	2	9	18
	問34	x													x		9	2	82
	問35			x				x		x					x	x	6	5	55
	問36					x											10	1	91
	問37																11	0	100
	問38					x									x		9	2	82
【多型解析】	問39															13	0	100	
	問40															13	0	100	
	問41															13	0	100	
	問42					x										12	1	92	
	問43	x		x		x	x	x	x					x	x	5	8	38	
	問44				x		x		x					x		9	4	69	
【プロテオーム解析】	問45															13	0	100	
	問46															13	0	100	
	問47									x						12	1	92	
	問48															13	0	100	
	問49														x		12	1	92
	問50	x				x	x	x							x		8	5	62
回答数		50	50	50	50	50	50	50	50	26	50	6	21	50	50				
正答		41	44	37	45	43	39	40	34	17	42	6	19	30	34				

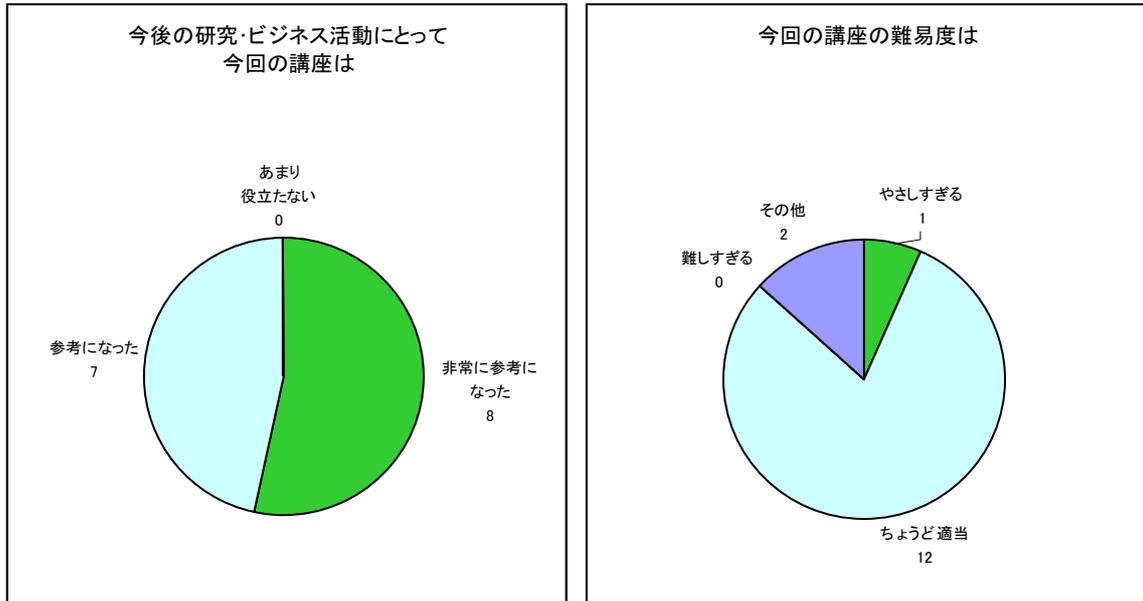
回答者/問	A	B	C	D	E	F	G	H	I	J	K	L	M	N	○	×	正答率 (%)
正答率 (%)	82	88	74	90	86	78	80	68	65	84	100	90	60	68			

#### D. 実証講座モニター受講者毎回アンケート

講座の受講者に対して、カリキュラムその他に意見を反映させることを目的として、毎回授業終了後にアンケートを実施した。設問は次の通りとし、毎回同じものとしたが、第5回においてのみ、追加での質問を行った。

- 1．今回の講義は、あなたの今後の研究・ビジネス活動にとって、(選択)  
非常に参考になった      参考になった      あまり役立たない
- 2．今回の講義の難易度はいかがでしたか。(選択)  
ちょうど適当      難しすぎる      やさしすぎる      その他(                      )
- 3．今回の講義のよかった点を記載してください。(自由記述)
- 4．今回の講義の改善すべき点を記載してください。(自由記述)
- 5．その他、ご意見を自由にお書きください(自由記述)
- 6．アノテーターとは(第5回のみ)(自由記述)

## No.1 配列解析 1



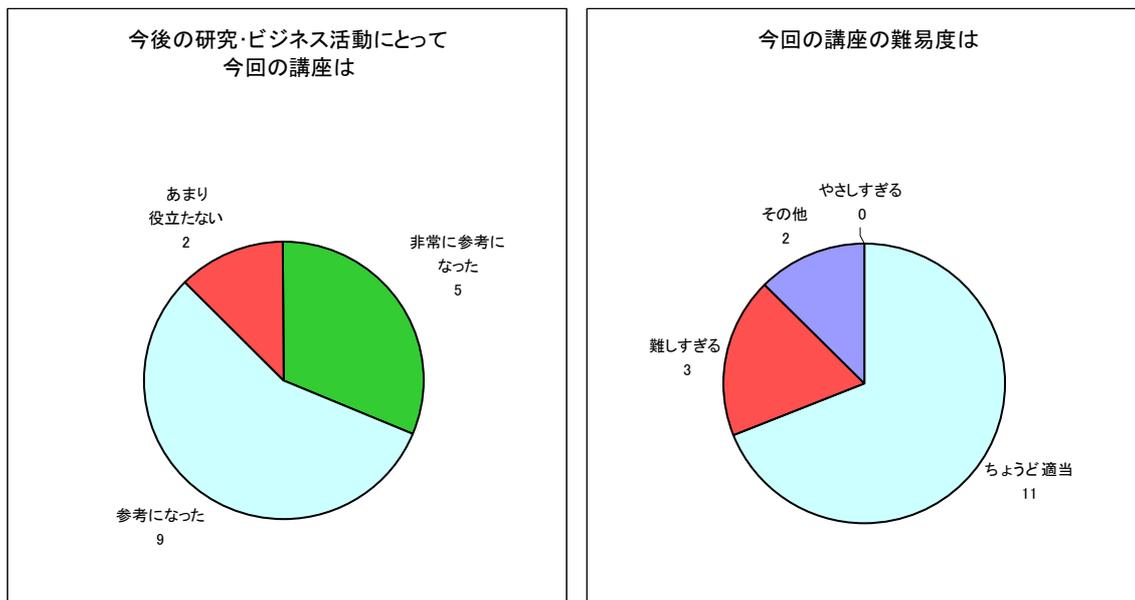
### 今回の講義のよかった点

- ・ 遺伝子の塩基配列からタンパクの配列、特徴の予想まで行う方法はなんとなくわかっていたのですが、実際にやったことがなかったので、実習付きというところがよかったと思います。
- ・ step wise で一步一步確認しながら、data base の使用法、応用について説明して下さいました。
- ・ 一連の流れが網羅されていて良かったと思う。
- ・ 実習を多く取り入れている点。
- ・ 実際に pc をさわりながら、講習を受けられた点。
- ・ 一通りの方法が学べたので、アノテーターの全体像がわかってきたように思います。
- ・ Link 集を作っていただき、入力に時間がとられなかった。

### 今回の講義の改善すべき点など

- ・ Bio informatics の総論的内容について触れて欲しかった。先生御自身の成功例を交えながらお話いただきたいと思います。
- ・ アノテーターとしての、使用頻度の高い順で講義したほうが、理解しやすいと思います。
- ・ 色々な DB での検索方法、調べ方を具体的に説明された点が良い。
- ・ 一度にアクセスが集中するせいか、結果が出るまでに少し時間がかかっていたように思いました。
- ・ 質問の時間をもう少しつくて欲しい。
- ・ 時間の配分を考えてほしい。
- ・ どのようなレベルの受講者を対象としていたのか、募集時から開始まではっきりしていなかったため、全員が同じことを感じたか疑問。
- ・ 予習・復習を受講生に求めることが講習計画の中に組み込まれていたほうが、講習としての成果と満足感が高いと思います。

## No.2 配列解析 2



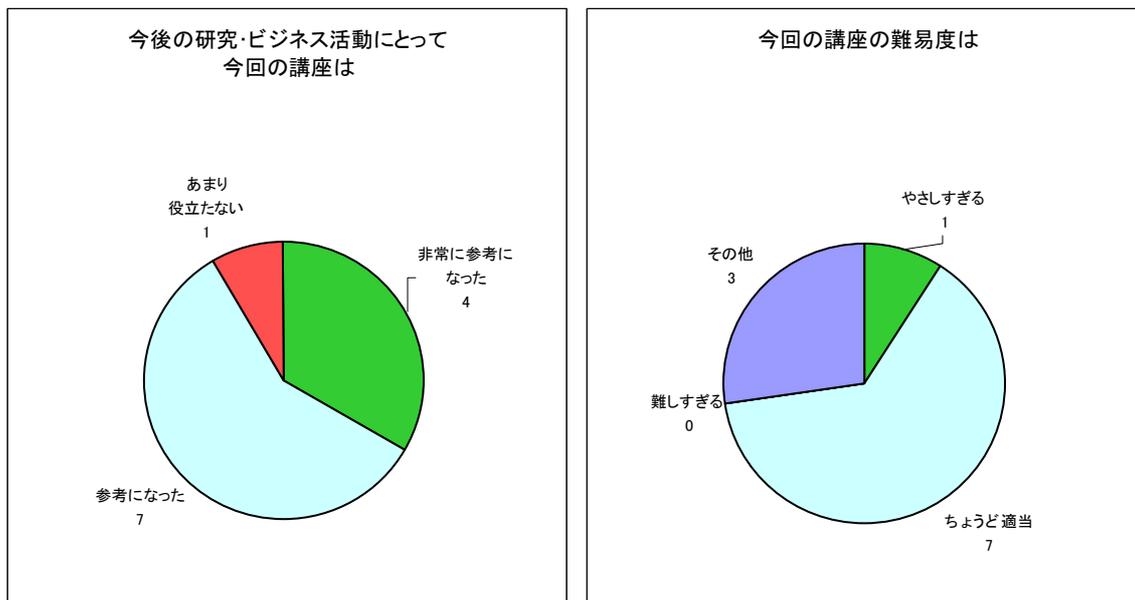
### 今回の講義のよかった点

- ・ 実際の流れをなぞる形であったのが良かった。一般的な知識の紹介より、実際に使うツールについて詳しく話されるのは為になります。
- ・ standalone での解析法を説明していただいた点は参考になりました。
- ・ CD ベースで既に教材が用意されていたのが良かった。
- ・ 演習問題用に、タイプの異なる配列データをたくさん用意していただいていたので、実際のアノテーションを類似体験することができた。
- ・ 日ごろの疑問点が解決。
- ・ 用語の説明もあり、大変よくわかりました。

### 今回の講義の改善すべき点、その他自由意見など

- ・ ビューアーArtemis の使い方が、最初とまどった。
- ・ バッチが組まれているのは簡単で良かった一方、何をしているのかを把握しづらい状況でした。一つずつ確認しながら進めていただければと思います。
- ・ 今回は、手順どおりに行えばよいと思いますが、応用が難しいと思います。
- ・ ところどころ説明があやふや、あるいはほとんど説明しないまま進行することがあった。
- ・ ペースが早いのではないかと思います。
- ・ PC が遅いです。
- ・ 前回の講義とのつながりがもう少し工夫されても良い。オムニバス形式の欠点。
- ・ かなりのバイオ知識を前提としているのかと感じた。
- ・ 自分としての必要とするレベルより、少し高いところに及んでおり、可能なら前もっての講義内容提示が詳しくなされるとよいのではないかと。

### No.3 配列解析 3



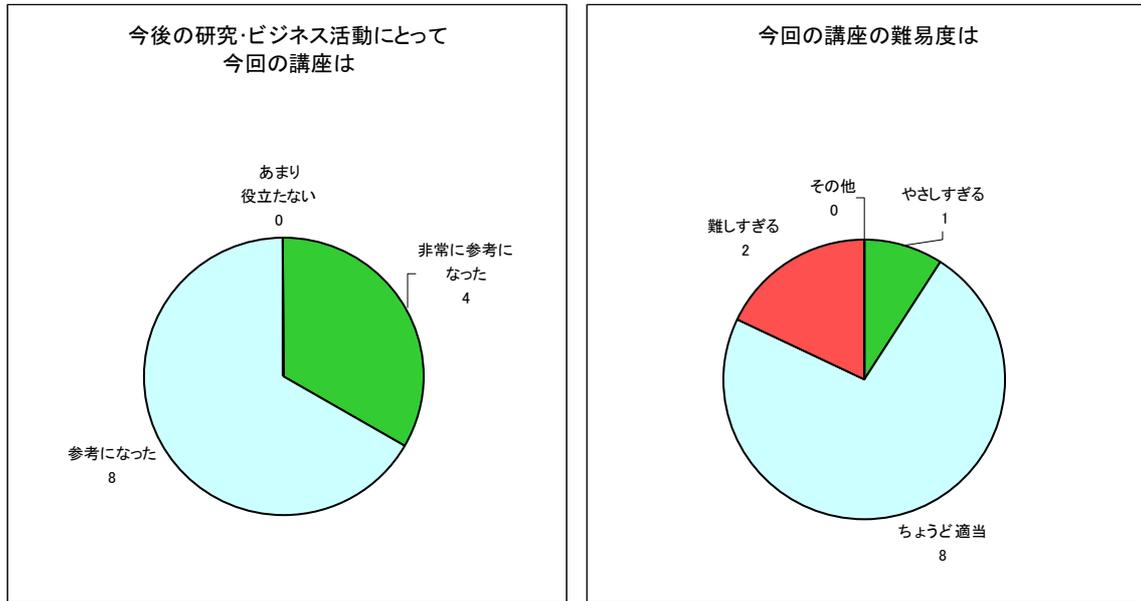
#### 今回の講義のよかった点

- ・ 導入に前回までのまとめを概説しており、今回の実習の目的を充分理解させていた。
- ・ 系統樹の解析に多くの時間を費やして下さり、専門外の者でも、その意味、作成の方法がよくわかった。
- ・ 進化・種間は知らない分野だったので実際にどのような形で機能理解に役立つのかが分かった。
- ・ 現場の話が出ていて、現場の声が伝わってくる講義だったと思う。
- ・ 話がわかりやすかった。説明も詳しくて良かった。声が良く聞きとれて良かった。

#### 今回の講義の改善すべき点、その他自由意見など

- ・ 時間配分やコンピューター環境など、もう少し事前準備をした方が良いと思う。
- ・ サイト等はプレインストールしておいて頂くと個々人の作業差が減り良いと考える。
- ・ 利用するプログラムのインストール手順で説明が速すぎた。
- ・ アノテーションを実際の研究でどう生かしてられるか成功例を挙げて説明して頂きたかった。
- ・ アドミニストレータの権限がないと、インストール出来ない。
- ・ URL やファイルが必要だったら配布したり自由に取れるディスクスペースがあればよいような気がする。
- ・ PCの事前作動チェックをおねがいします。

## No.4 発現プロファイル解析 1



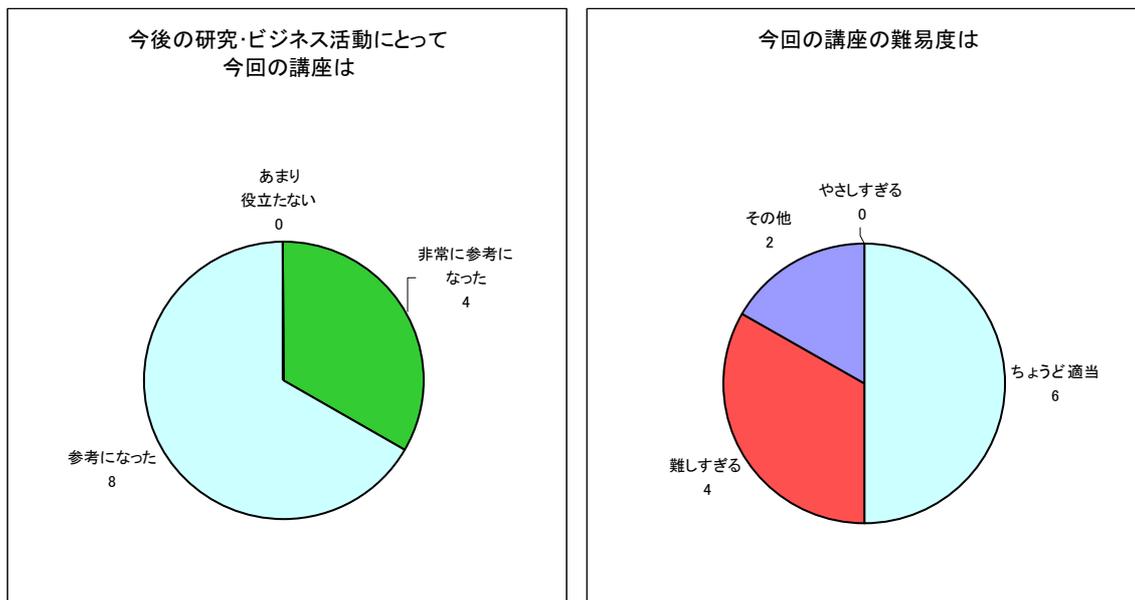
### 今回の講義のよかった点

- ・ 一人で3回の講義を担当するため、全体の概要説明が充分に行なわれ取り組みやすかった。
- ・ アレイの基礎的な解析を説明された点。
- ・ マイクロアレイの利点と短所が理解できた。
- ・ システムバイオロジーの基礎的な事からの丁寧な説明であった点。
- ・ Rに触れたこと。
- ・ 各 Method の pnos/cons を詳細に説明して下さったこと。

### 今回の講義の改善すべき点、その他自由意見など

- ・ 演習と講義の時間の配分について、可能なら演習を重点的にしていただければと思います。
- ・ もっと「R」をつかえると良かった。
- ・ 色々なプログラム、コマンドのリストが欲しいです。
- ・ 実験の理論と解析の理論をもっと丁寧に説明してほしい。データの解析方法は分かったのだが、そのデータがどう出てきたかがまだイメージできず、そのために理解度が低い。
- ・ 疾患と結びつけて、先生の経験あるいは過去の成功例を特殊な terminology の説明をもっとくわしく提示してほしい。
- ・ ちょっと生物系の方から見ると、難しい内容でした。
- ・ 参加者のレベルがばらばらなのではない部分はある。

## No.5 発現プロファイル解析 2



### 今回の講義のよかった点

- ・ 解析の流れに沿った講義で、実践的な内容であった。
- ・ R という統計ソフトのすごさの一端をかいま見た。
- ・ 演習中心で解析の実際を経験できて有意義でした。
- ・ 実際に入力することができたので、細かい疑問点がある場で理解できた。

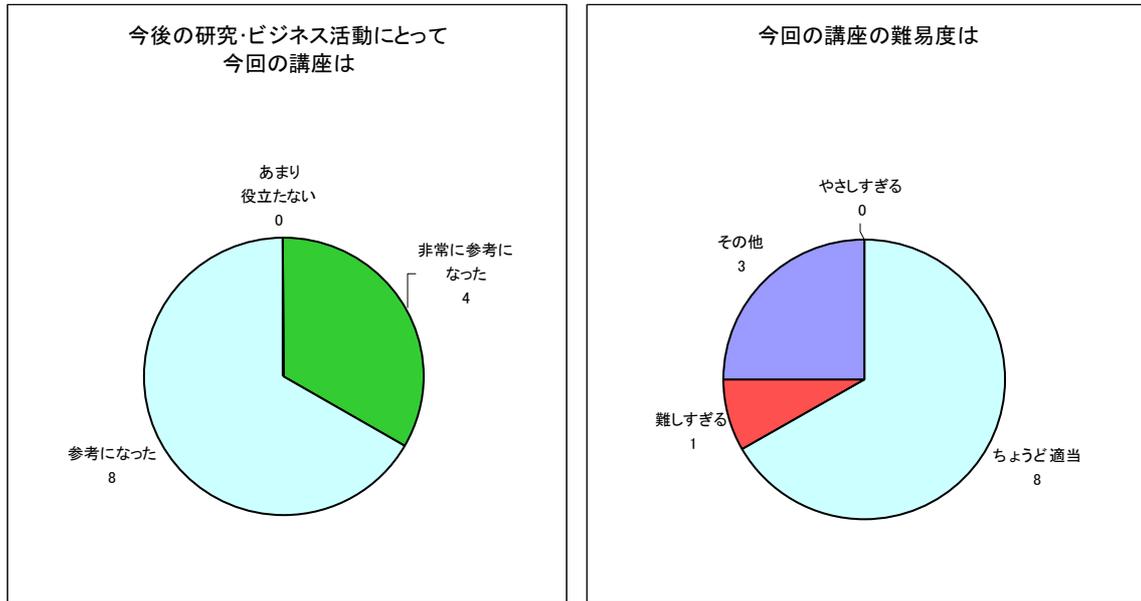
### 今回の講義の改善すべき点、その他自由意見など

- ・ プログラムの一つ一つの意味を簡潔に説明してください。
- ・ コマンドを打ち込むだけで精一杯で、何をやっているのかわからなかった。
- ・ "R" の使用方法に終始したきらいがあり、本来の発現プロファイルの data の解釈についての説明が乏しい。
- ・ コマンドラインによる操作を行なうものは、受講者に行なわせるというより、理解させることが重要。数学的なものは空々しい。詳しい説明が必要かもしれない。
- ・ ひとつひとつの操作はどういう原理でどういう意味を持つのか教えてほしい。

### アノテーターとは

- ・ 遺伝子に意味をつける人のこと。意味づけは配列の特徴はもちろん発現についての情報など可能な限り付与する。
- ・ Wet の実験と dry の data の link を行い、より少ない自験データから最大限の効果的情報を得る専門家。
- ・ 事象に対して意味説明をつける作業。Gene に限らず生物の形状、形態、化学物質、化学反応、それらからなる現象に対する説明付けも必要になってくると思われる。
- ・ アノテーションを専門とするスペシャリストとだけ捕らえるのは難しい。ゲノミクスからメタボロミクスまで、総合的な知識と技能を必要とするジェネラリスト的な人材と考える。
- ・ Wet 研究者が実験で出したバイオデータを IT で解析することで、IT 企業が構築した解析システムと Wet 研究者との 結びつける人材がアノテーター。
- ・ 新規遺伝子を見つけた人がアノテートするものとは思っている。やはり、今後ジーンチップなどハイスループットのデータが大量にできた時に職業として確立していくのだろう。
- ・ 解析を考える手がかりになるアノテーションを集めて、整形する人。
- ・ 実験から得た結果を多面的に解釈する手段。

## No.6 発現プロファイル解析 3



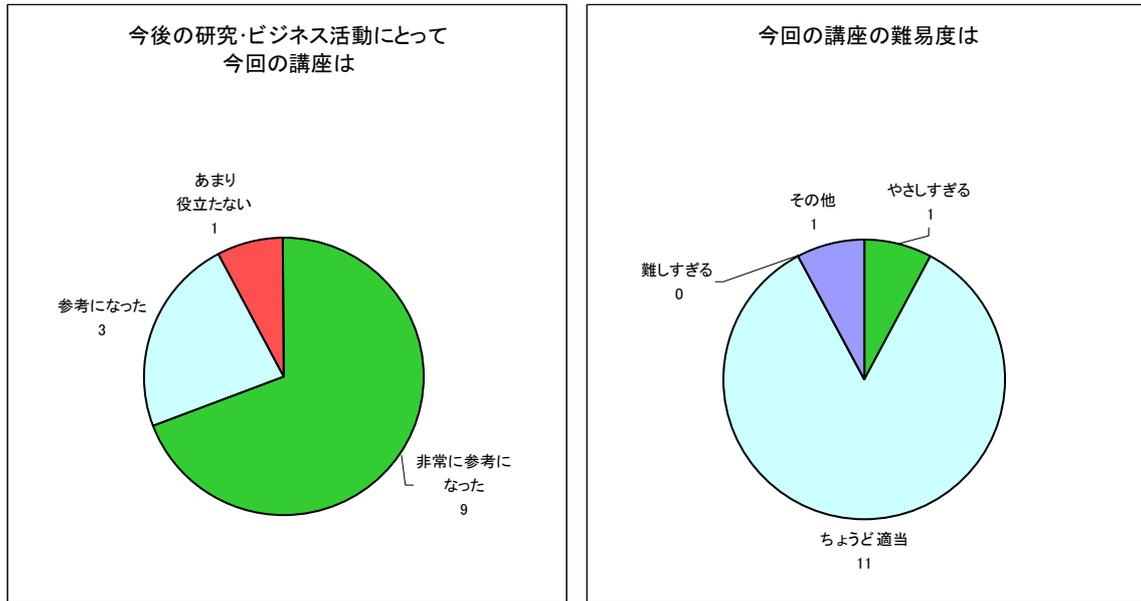
### 今回の講義のよかった点

- ・ Simulation プログラムをはじめて使い、Simulation の雰囲気味わうことができた。
- ・ 今回は入力作業が少なかったので良かったと思います。
- ・ プログラムがコピー＆ペーストでよかった。
- ・ シュミレーターソフトの演習は有用に思えた。

### 今回の講義の改善すべき点、その他自由意見など

- ・ 実際の解析結果として表示されたものをどう解釈するか、時間を割いたほうが良い。
- ・ R の本や R tips の紹介があると、初心者親切だったかも。
- ・ どうしてこのような分析を行なうのか？とかもっと詳しい説明がほしかったです。それから、シミュレーションやネットワーク推定とかもう少し説明してほしかったです。
- ・ 実験データを元にコンピューターを使って生物学的な意味を見出すのがアナテーターであるはずなのに、生物学的な解釈の話が少なすぎる。「マイクロアレイ ネットワーク推定 シミュレーション」の一連の流れの概要はわかったが、実際にどうすればよいかは今回の講義では理解できない。
- ・ 内容が浅く広くなってしまい、何を説明したいのかわかみにくい。
- ・ 現実の解析で、どのようなツールを使って解析しているか、具体的に知りたい。
- ・ 3時間×3回ではなく、もっとゆっくり聴きたかった。

## No.7 多型解析 (SNP)



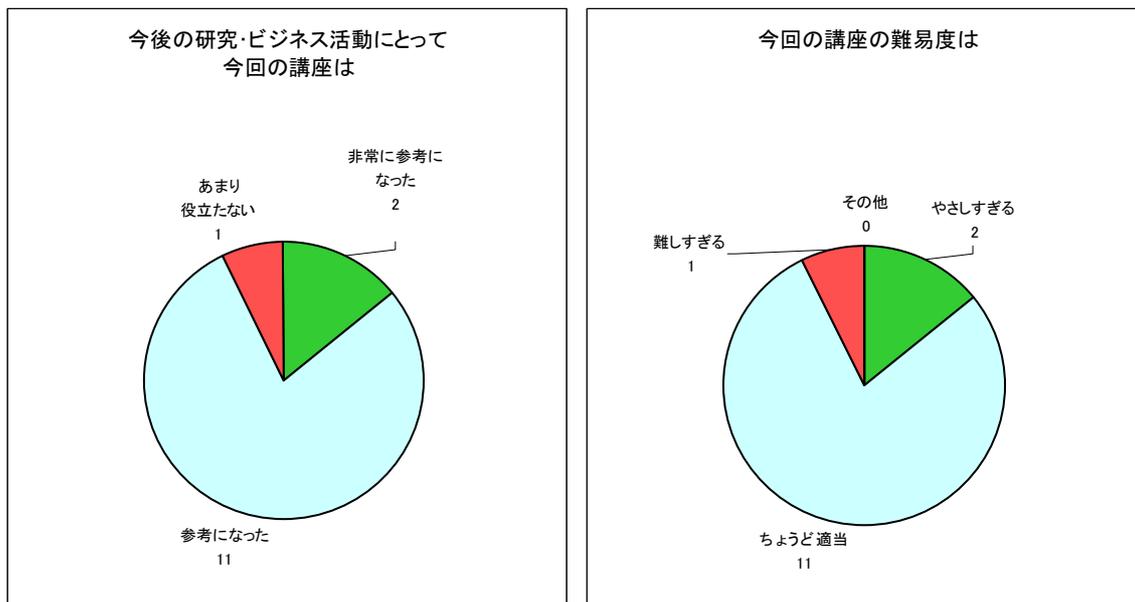
### 今回の講義のよかった点

- ・ ヒトゲノムの多様性について総括できた。SNP 全般について over view できた。
- ・ 講義の内容が良く整理されていてわかりやすかった。ポイントをおさえた内容で良かった。
- ・ 進行速度も良く、バランスが取れていた。課題の作業時間も確認を取りながら丁度よかった。
- ・ 基本から順に説明されましたし、演習もきちんと分かるようで良かったです。
- ・ 統計(計算)の部分があったこと。

### 今回の講義の改善すべき点、その他自由意見など

- ・ 演習は問題、例題を書いた用紙がないとついて行くのがとても困難になる。
- ・ Hand out が無かったので、筆記が大変。
- ・ db の使用法の詳細 (eg . db SNP の画面コピーなど) も資料に載せてください。強く希望します。
- ・ もう少しフリーソフト等を使った演習が良かった。
- ・ Database の利用の仕方をもう少し知りたかった。
- ・ 祖先型 SNP の話がとても面白かったです。
- ・ 応用編として実際の統計解析をもう一回分設定するとよい。
- ・ 資料は可能な限り当日頂きたいです。

## No.8 プロテオーム解析とタンパク質相互解析



### 今回の講義のよかった点

- ・ 工学系というよりは、実際に実験を行っている先生に説明していただいたのは良かったです。
- ・ Protein 関連 Database を多数紹介くださった点は良かった。
- ・ 未知タンパクの具体的解析手法としても興味深かった。
- ・ プロテオーム解析に必要なデータベースの利用法等。
- ・ 新しいDB (MINT) を知ったこと。
- ・ mascot 検索を自分で行ったことはなかったなので、実際やってみることで理解が深まった。

### 今回の講義の改善すべき点、その他自由意見など

- ・ データベース以外で解析可能なソフトの説明等を加えていただければよかったのではないかと思います。実験の説明よりバイオインフォマティクスの部分をより多く取り入れていただければと考えます。
- ・ yeast や bacteria ではなく、human のタンパク相互作用を解析できる HP を紹介してほしいです。あともう少し転写のこともやっていただけたらよかったです。
- ・ 先生は実際に実験をなさっている方とのことですので、ご自身の実データを中心に説明していただきたかった。
- ・ 演習につながりがあると良かったです。例えば MASCOT で複数のたんぱく質を同定し、その後 BIND 当の相互作用 DB で、それらのたんぱく質間に関係があるのかどうかを見るなど。
- ・ 発現解析の演習ではもう少し応用的な検索法もやってほしいです。
- ・ mascot の結果を詳しく説明するほうが良い。
- ・ 2 回に分けて MS のデータ処理、基本的な知識とデータベース検索結果の解釈 相互作用など にしないと、内容が浅くなりすぎ。

## E. 実証講座モニター受講者事前アンケート

実証講座を開催するにあたり、受講者に対し、受講内容に対する理解度を自己評価させた。設問は配列解析、発現プロファイル解析、多型解析 (SNP)、プロテオーム解析、タンパク質間ネットワーク解析の5つの実験分野について、「実験の知識・理解」「解析ツールの活用」「解析結果の解釈」という能力ごとに、自身について「よく理解している」「概略は理解している」「よくわかっていない」の三者択一の評価をすることとした。

回答は、モニターの15名全員から得た。

		よく理解している		概略は理解している		よくわかっていない	
配列解析	実験の知識・理解	3	(20%)	11	(73%)	1	(7%)
	解析ツールの活用	2	(13%)	8	(53%)	5	(33%)
	解析結果の解釈	2	(13%)	8	(53%)	5	(33%)
	(平均)		(15.6%)		(60.0%)		(24.4%)
発現プロファイル解析	実験の知識・理解	3	(20%)	10	(67%)	2	(13%)
	解析ツールの活用	1	(7%)	5	(33%)	9	(60%)
	解析結果の解釈	1	(7%)	6	(40%)	8	(53%)
	(平均)		(11.1%)		(46.7%)		(42.2%)
多型解析 (SNP)	実験の知識・理解	2	(13%)	9	(60%)	4	(27%)
	解析ツールの活用	1	(7%)	5	(33%)	9	(60%)
	解析結果の解釈	1	(7%)	6	(40%)	8	(53%)
	(平均)		(8.9%)		(44.4%)		(46.7%)
プロテオーム解析	実験の知識・理解	2	(13%)	7	(47%)	6	(40%)
	解析ツールの活用	1	(7%)	5	(33%)	9	(60%)
	解析結果の解釈	1	(7%)	6	(40%)	8	(53%)
	(平均)		(8.9%)		(40.0%)		(51.1%)
タンパク質間ネットワーク解析	実験の知識・理解	2	(13%)	8	(53%)	5	(33%)
	解析ツールの活用	1	(7%)	5	(33%)	9	(60%)
	解析結果の解釈	1	(7%)	4	(27%)	10	(67%)
	(平均)		(8.9%)		(37.8%)		(53.3%)

## F. 実証講座モニター受講者事後アンケート

実証講座のすべての授業が終了した後、理解度や実証講座における問題点、改善点等を把握するため、モニター受講者に対して質問した。質問・回答項目は事前アンケート同様に、「各実験」についての「実験に関する能力」を、受講前と比べて「よく理解できるようになった」「概略は理解できるようになった」「あまり変わらない」の三者択一によることとした。また、あわせて本講座全体を通しての質問と回答を以下にまとめた。

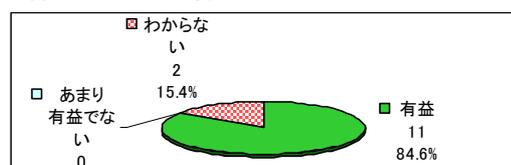
なお、回答は13名より得ることができた。

<受講前と比較して>		よく理解できるようになった		概略は理解できるようになった		あまり変わらない	
配列解析	実験の知識・理解	4	(31%)	9	(69%)	0	(0%)
	解析ツールの活用	4	(31%)	9	(69%)	0	(0%)
	解析結果の解釈	2	(15%)	11	(85%)	0	(0%)
	(平均)		(22.2%)		(64.4%)		(0.0%)
発現プロファイル解析	実験の知識・理解	2	(15%)	8	(62%)	2	(15%)
	解析ツールの活用	1	(8%)	9	(69%)	2	(15%)
	解析結果の解釈	1	(8%)	8	(62%)	3	(23%)
	(平均)		(8.9%)		(55.6%)		(15.6%)
多型解析 (SNP)	実験の知識・理解	5	(38%)	6	(46%)	1	(8%)
	解析ツールの活用	2	(15%)	7	(54%)	3	(23%)
	解析結果の解釈	3	(23%)	5	(38%)	4	(31%)
	(平均)		(22.2%)		(40.0%)		(17.8%)
プロテオーム解析	実験の知識・理解	2	(15%)	9	(69%)	2	(15%)
	解析ツールの活用	2	(15%)	10	(77%)	1	(8%)
	解析結果の解釈	2	(15%)	8	(62%)	3	(23%)
	(平均)		(13.3%)		(60.0%)		(13.3%)
タンパク質間ネットワーク解析	実験の知識・理解	1	(8%)	12	(92%)	0	(0%)
	解析ツールの活用	0	(0%)	13	(100%)	0	(0%)
	解析結果の解釈	1	(8%)	10	(77%)	2	(15%)
	(平均)		(4.4%)		(77.8%)		(4.4%)

[ 本講座全体を通じて ]

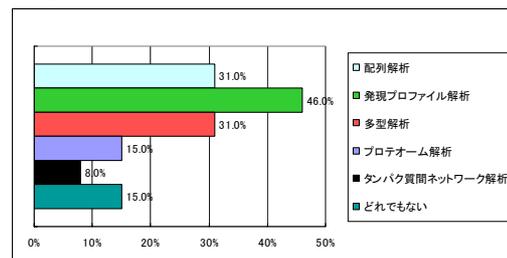
( 1 ) 本講座は受講者自身の今後の研究・ビジネス活動にとって有益でしたか

有益	11	( 84.6% )
あまり有益でない	0	( 0.0% )
わからない	2	( 15.4% )



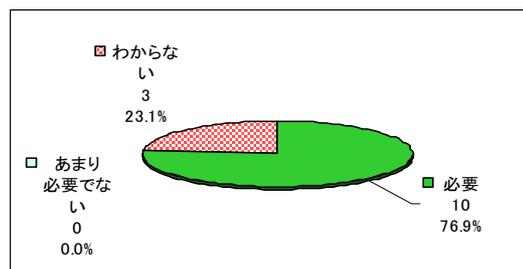
(2) 最も有益であった講義分野は(複数回答)(N=13)

配列解析	4	(31.0%)
発現プロファイル解析	6	(46.0%)
多型解析	4	(31.0%)
プロテオーム解析	2	(15.0%)
タンパク質間ネットワーク解析	1	(8.0%)
どれもでない	2	(15.0%)



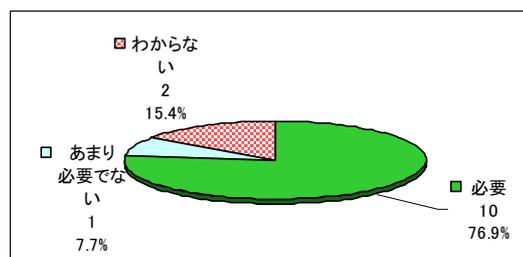
(3) アノテーター人材の必要性

必要	10	(76.9%)
あまり必要でない	0	(0.0%)
わからない	3	(23.1%)



(4) 今後の研修の必要性

必要	10	(76.9%)
あまり必要でない	1	(7.7%)
わからない	2	(15.4%)



(5) 研修を希望されるバイオインフォマティクス、アノテーションの分野

- ・ DBの相互利用
- ・ 配列、発現解析、多型解析
- ・ SNP, Proteome
- ・ 2 でつけたところ (注: 配列解析、発現プロファイル解析、プロテオーム解析、タンパク質間ネットワーク解析)
- ・ 疾患と分子の関連
- ・ 統計

(6) 受講者の経歴について

バイオ経験年数		IT 経験年数	
~1年	1	~1年	4
~5年	2	~5年	3
~10年	3	~10年	1
~15年	5	~15年	1
~20年	1	~20年	0
20年~	1	20年~	0

[主な経歴]

バイオインフォマティクス(4)、教員(2)、SNP、疾患関連遺伝子探索、遺伝子発現制御メカニズム解析、組み換えニワトリ用のワクチンの開発、ウイルス改変、外来遺伝子導入、創薬、二次代謝産物の生合成酵素の構造、機能解析、バイオプロセス関連のサポート業務

(7) よく使うデータベース

NCBI (6)、Ensembl(2)、JSNP(2)、その他 (10)

(8) 開催日時やクラス、環境等についての受講者の意見

- ・ プロジェクターは大きいほうがよい。ほかはよいと思う。
- ・ 休日・昼、1クラス12～15人、教室はもう少し広いほうが、パソコンのスピードは充分、教室環境 OK.
- ・ ちょうどよいのでは
- ・ 今回のようなものが最適。時間はもう少し長くてよい
- ・ PC のメモリ以外は不満なし。初回としては満点に近い。
- ・ パソコンの性能にやや課題あり
- ・ パソコンでソフトを使う際に制限があったのは少し改善していただけるとありがたいです
- ・ 平日の夜のほうがよかったです。
- ・ 開催日時は平日(特に水曜)昼間がよかった。1クラスの人数は20人くらい、教室はもう少し広いほうがよい。PC は能力的に問題は無かったが、admin が必要なソフトへの対応が不十分。Web 検索では、同一 IP からのアクセスで制限がかかることが多いのでそれに対する対応も必要
- ・ おおむね OK
- ・ 今回のような予定で問題ないと思います。
- ・ 特になし
- ・ 一週間に半日を費やすのは難しいです。2～3時間ずつで、倍の期間であればよいと思います

(9) その他の意見自由記述

- ・ 予習・復習用の教材を用意したほうがよい。
- ・ バイオインフォマティクスは初めてだったのですが、参考になりました。学生の教育で生かされるとよいと思います。
- ・ 多々、参考になりました。参加出来てよかったと思います。
- ・ 講師の先生方によっては時間配分がよい先生とそうでない先生とがハッキリしていました。発現プロファイルの三回は、入力する作業(2回目)スライドがどんどん進むのみ(3回目)というだけで、内容(ex,このコマンドは何を意味するのか)をご説明いただきかったです。実習中心ということでしたが、もっと内容を説明する、言い換えれば各実習で行なっている作業内容(SNP の講師の先生はよく理解できました)を説明してほしいです。
- ・ モニター応募資格をもう少し限定したほうがよいと思います。又、解析を行なう前に行なう実験手法についても簡単に説明してもらえたらよかったと思います。解析中心であまり内容を理解できなかった部分もあったので、解析方法についてももう少し詳しく説明していただけたらと思いました。
- ・ Power Point で講義をしているのだからそのファイルをどこかのサーバーか FTP において生徒がそのファイルを取りにいく(見に行く)ことによって、入力間違いでの時間のロスはなくなると思われる。
- ・ 充実名内容でした。ありがとうございました。まだ自身の知識が不足していたため、難しかったですが、非常に有益な講義でした。
- ・ 長期出張が予想外に重なったために出席できず、大変ご迷惑をおかけしました。申し訳ございませんでした。