

人材の育成

V. 人材の育成（長期的な取り組みにむけて）



データベース統合や維持管理のためには、キュレーター(論文の内容を理解して情報抽出整理し定型化した表現に変換してDB構築を支援する学芸員)、アノテーター(プログラム処理の結果に総合判断を加えてデータに生物学的医学的な解釈を付与する学芸員)、DBマネージャー(DBについて理解しておりデータの参照情報などを自立的に更新できる技術者)の三者が必要です。しかしながらこれらの専門業務の存在はデータベースの構築運営を行ってきた組織でしか知られていません。広くこれらの業務内容について衆知し、学習の材料を与えることは将来のデータベース統合事業に参加可能な人材の裾野を広げることであり、これらの専門職のキャリアパス作成の第一段階であると考えられます。ここでは三者の業務内容を整理し、業務に必要な基礎知識や技術について解説した教材を作成してデータベースの本格的統合化事業に備える目的で以下の事業を実施しました。

平成18年度は、学習のための教科書を編纂、提供するためのシステムとして、wikiと電子メールによって知識の蓄積が可能な教育資料編纂・閲覧システム「MotDB」を qwikweb をベースに作成しました。

教科書編纂・提供システムに qwikweb を利用したことで、wikiによるWWWからの追記だけではなく、電子メールからの入力も可能となっています。これにより、より多くの有識者から情報を集積し、将来にわたり継続的に情報を更新できる体制をとることを可能にしました。

(1) キュレーター教育

ハイスループット化されたゲノム研究手法により、大量のゲノム配列情報が作り出され、膨大な生物種のゲノム情報が刻々と蓄積しつづけています。同時に、従来の小規模で精密な実験やゲノム規模の実験による、発現プロファイル解析、タンパク質相互作用解析や構造ゲノミクス解析、また、それに伴う計算機での大規模予測がなされ、さまざまな角度からタンパク質機能の解析がなされています。それらのデータの統合活用には、自動アノテーションのチェックや文献からの情報抽出やコンピュータ解析によるアノテーションの修正と付加の過程であるキュレーションの整備と、その担い手であるキュレーターの育成が必要かつ急務であると考えられます。

このような背景から、キュレーションおよびキュレーターの教育を行うための基礎知識や技術についての教材を作成するための調査を実施しました。

調査は平成18年度統合データベース基盤整備プロジェクト研究員(II)岡本忍が論文とインターネット上から得られる情報を元に実施しました。また、同プロジェクトの「人材の育成」担当グループから構成される「MotDB(アノテーター、キュレーター、システムデータベース管理者向けの教材を構築していくWikiサイ

ト)」(代表:坊農秀雅助教授、埼玉医科大学ゲノム医学センター)にて調査手法や分析結果に対するご意見および、課題等の提案をいただきました。

生物の基本的かつ包括的な情報を扱う代表的なキュレーション型DBとして、RefSeq (NCBI), UniProtKB (EBI, PIR, SIB), KEGG PATHWAY (BiC)について調査しました。各々のDBの特徴と規模、どのような体制でキュレーションが行われているか、またアノテーション過程とキュレーション過程の関係や、キュレーターが担う実際の仕事の範囲について検討しました。さらに人材としてキュレーターに求められる基礎知識や技術、資質についても調査検討しました。

(2) アノテーター教育

アノテーター教育用の支援プログラムと教科書を作成しています。まず、これまで数多くの高度ゲノムアノテーションを実施してきたかずさDNA研究所植物ゲノム基盤研究部(※1)のアノテーターが持つ世界最高レベルのゲノムアノテーション技法について、聴き取りと実務の再現による詳細なゲノム解析手順の調査によりノウハウを整理し、その手順の大枠をプログラム化して判断根拠とともに明示する形でアノテーター教育に役立つシステムの基盤部分を作成しました。この調査の結果を参照しつつ、生物学者に利用しやすい形でゲノム上の特徴情報を付加するために必要となる解析法の基礎を含んだ「アノテーション」教育用日本語テキスト作成を試みています。微生物ゲノムの配列解析に関しては、平成18年度、長浜バイオ大学の三年生のゲノムアノテーション実習と、四年生の卒業研究のテーマとして、まったくの初学者が新規微生物のゲノム塩基配列の大規模解析にトライしたプロジェクトの結果も参考にしています。

※1:かずさDNA研究所は、1996年のシアノバクテリア(酸素発生型光合成細菌)Synechocystis PCC 6803での、日本初の微生物全ゲノム塩基配列解析を始めとして、2000年には世界初の高等植物の全ゲノム塩基配列解析の報告となったシロイヌナズナゲノムプロジェクト(国際コンソーシアム)に参加し、全ゲノム115 Mbの約1/4の配列決定とその領域の一次アノテーションを担当するなど、多くのゲノム解析プロジェクトを行ってきています。微生物ならびに高等植物分野で定評のある解析を実現した、大規模ゲノム塩基配列解析のためのアノテーションパイプラインとアノテーターを擁しています。

(3) DBマネージャー教育

DBマネージャー教育用の教科書を作成しています。DBを管理するDBマネージャーに必要な不可欠なスキルとして大きくDB構築とその維持がありますので、教科書もそれに合わせて大きく、構築編と維持管理編の二つに分け、平成18年度は既存のDBの維持管理に必要なスキルを解説する維持管理編の構築を主に進めました。その過程でこの二つのカテゴリーには入らない事柄を、DBマネージャーとして前提とする基礎的な知識を示した基礎編、さらに各編に共通したリファレンスとなりうる事項を蓄えるリファレンス編としました。

構築した維持管理編は、日々の管理、アップデートとバックアップに分け、「DASサーバーの維持管理」を具体例として必要なコマンド操作を解説しています。