



IMSUT & RCAST GCOE 特別セミナー 〈キャリアパス支援テクニカルセミナー〉

「HGCスパコン Web サービス利用法講習会」開催のご案内

(医科研ヒトゲノム解析センター・スーパーコンピュータ室 主催)

この度「HGC スパコン Web サービス利用法講習会」と題しまして、ヒトゲノム解析センターで利用可能な Web サービスを紹介致します。日程と内容は下記の通りです。興味のある部分だけの参加も可能です。広く皆様のご参加をお待ちしています。

日時：2011 年 10 月 25 日(火) 10:00 - 17:00

会場：医科学研究所ヒトゲノム解析センター 4 階セミナー室

定員：34 名（事前申込者を優先、空席があれば当日参加も可能）

参加費：無料

* 本講習会は日本語による講義です（一部英語です）。

10:00 - 10:30 HGC スパコンサービスの紹介

10:30 - 11:00 GeneCards の紹介

11:00 - 11:30 真核生物転写制御解析のためのデータベース：DBTSS と Melina2

11:30 - 12:00 共発現データベース COXPRESdb

12:00 - 13:00 休憩

13:00 - 13:30 枯草菌転写データベース DBTBS と細胞内局在部位予測サーバ PSORT

13:30 - 14:00 Comprehensive and reliable protein-protein interactions using HitPredict and Hintdb

14:00 - 14:30 PDBj & eF-site の紹介

14:30 - 14:40 休憩

14:40 - 15:10 KEGG データベースの紹介

15:10 - 15:40 ライフサイエンス統合 DB のサービス紹介

15:40 - 16:00 休憩

16:00 - 16:30 Cell Illustrator Online の紹介

16:30 - 17:00 大規模遺伝子ネットワーク推定ソフトウェア SiGN

[HGC スパコンサービスの紹介] システムエンジニア

HGC のスーパーコンピュータシステムは、ゲノム分野では国内最大級の計算機システムで、各種サーバや高速大容量のディスクアレイ装置によって多様なゲノム研究環境を提供している。講習会では、HGC スーパーコンピュータシステムの概要、有償利用サービスの概要について紹介する。

[GeneCards の紹介] 小池麻子 ((株)日立製作所中央研究所)

GeneCards は Weizmann Institute で構築されている統合型データベースである。

遺伝子の配列情報だけでなく、蛋白質相互作用、遺伝子機能情報、SNP 情報化合物情報など様々な情報を一度に検索することができる。GeneCards に含まれる情報と効率的な検索方法について解説する。

[真確生物転写制御解析のためのデータベース : DBTSS と Melina2]

山下理宇 (東京大学医科学研究所ヒトゲノム解析センター機能解析イン・シリコ分野特任助教)

DBTSS は、ヒト・マウス等の転写開始点データベースである。今年度このデータベースの大幅なアップデートを行い、ヒストン修飾や RNA ポリメラーゼの情報を追加した。

講習会では、DBTSS から目的の遺伝子の転写開始点情報を得る方法を紹介する。また、転写開始点付近の SNP 情報や転写制御領域の解析方法について紹介する。

さらに Melina2 を用いた複数の配列に潜む未知モチーフの発見法について紹介する。

[共発現データベース COXPRESdb] 大林武 (東北大学大学院情報科学研究科応用情報科学専攻助教)

COXPRESdb は 1000 以上の公共のマイクロアレイデータから計算したヒト・マウス・ラット・ニワトリ・ゼブラフィッシュ・ショウジョウバエ・線虫の共発現遺伝子データベースである。

各生物種について約 20000 遺伝子の共発現情報を検索することが可能であり、これら共発現遺伝子は機能的関連を持つ遺伝子であると期待できる。共発現情報の種間比較を行うことで、信頼性の高い機能推定が可能である。

また、自前のマイクロアレイデータ等の大規模データの解析に利用することも可能であり、講習会では GeneChip を用いた実験を行った後どのように COXPRESdb を利用するかについても紹介する。

[枯草菌転写データベース DBTBS と細胞内局在部位予測サーバ PSORT]

中井謙太 (東京大学医科学研究所ヒトゲノム解析センター機能解析イン・シリコ分野教授)

DBTBS は、グラム陽性菌モデル生物である枯草菌 (*Bacillus subtilis*) の転写関連データを集めたデータベースである。データベース中のデータは全て文献から集められ、転写の制御関係、転写開始点、終結点、オペロンなど転写に関連するデータを幅広く集めている。

一方、PSORT は細胞内局在部位予測プログラムであり、予測したいタンパク質の前駆体アミノ酸配列とそのタンパク質の由来生物の分類名を入力すれば、その配列中に存在する既知のソーティングシグナルの存在を検出し、アミノ酸組成等のその他の情報と組み合わせることで、可能性の高い局在部位を予測する。

[Comprehensive and reliable protein-protein interactions using HitPredict and Hintdb]

Ashwini Patil (東京大学医科学研究所ヒトゲノム解析センター機能解析イン・シリコ分野助教)

With the availability of a large number of protein-protein interactions,

it is imperative to assess their reliability prior to their use in further scientific studies.

HitPredict is a database that provides quality-assessed interactions in 9 species. Interactions are assigned with a score based on the genomic evidence found in their support. Such evidence can be obtained in the form of homologous interactions across species. The database, Hintdb, provides such a list of homologous interactions for those already determined experimentally. Together, HitPredict and Hintdb provide a useful resource of protein-protein interactions and their properties.

[PDBj & eF-site の紹介] 木下賢吾（東北大学大学院情報科学研究科応用情報科学専攻教授）
蛋白質立体構造情報は蛋白質の機能を理解する上で非常に有用な情報である。近年、たんぱく 3000 に代表されるような構造ゲノム科学プロジェクトの進展により、大量の立体構造情報が利用できるようになってきた。
本講習では、蛋白質立体構造情報の利用法の入門編として、立体構造情報の一次データベースである Protein Data Bank の日本支部（PDBj）で利用できるサービスの紹介と立体構造から計算できる分子表面と静電ポテンシャルのデータベース eF-site の紹介を行う。

[KEGG データベースの紹介]

川島秀一（東京大学医科学研究所ヒトゲノム解析センターゲノムデータベース分野助教）
KEGG (Kyoto Encyclopedia Genes and Genomes) は、京都大学化学研究所および東京大学医科学研究所で開発されている、ゲノム配列の決定された生物の遺伝子カタログ、生体内化合物、細胞内パスウェイ情報等を中心とした統合データベースである。複数のデータベースがどのように統合されているのかを中心に、基本的な検索方法および、解析に役立つツール群について解説する。

[ライフサイエンス統合 DB のサービス紹介]

片山俊明（東京大学医科学研究所ヒトゲノム解析センターゲノムデータベース分野助教）
ライフサイエンス統合データベースセンター（DBCLS）では、科学技術振興機構（JST）に今年度から設立されたバイオサイエンスデータベースセンター（NBDC）の統括のもと、様々な生命科学データベースを統合するための新しい技術開発に取り組んでいます。
データベースの横断検索をはじめ、文献検索、統合 TV、TogoDB、TogoWS などこれまでに開発されたサービスとともに、今後データの意味にもとづく統合化を担うとされる RDF 技術や国際開発者会議 BioHackathon の成果についてご紹介します。

[Cell Illustrator Online の紹介]

斉藤あゆむ（東京大学医科学研究所ヒトゲノム解析センターDNA 情報解析分野技術職員）
生体内パスウェイを描く機能とシミュレーションの機能が一体となっているソフトウェアです。CIO は、「生物系の研究室に所属する研究者が実験をするかたわら、パスウェイの情報を知識整理し、その知識整理したモデルをシミュレーションすることで次の実験に何をするかを決められる」ことを目標として開発しています。
講習会を通して、CIO を用いて基本的な生体内反応のモデル化の方法を学習するとともに、サーカディアンなどのモデルのシミュレーションを行う。

[大規模遺伝子ネットワーク推定ソフトウェア SiGN]

玉田嘉紀（東京大学大学院情報理工学系研究科コンピュータ科学専攻助教）

SiGN（サイン）は HGC のスーパーコンピュータを用いることにより、マイクロアレイなどで得られる遺伝子発現データから遺伝子間の発現の依存関係を表す遺伝子ネットワークを推定することのできるソフトウェアです。ベイジアンネットワーク・状態空間モデル・グラフィカルガウシアンモデル・ベクトル自己回帰モデルが利用可能（一部実装中）で遺伝子ノックダウン実験データ、患者サンプル由来のデータや薬剤投与時系列データなどを利用して、静的・動的な遺伝子ネットワークを推定することが可能です。

講習会では各遺伝子ネットワークモデルおよび推定法の紹介を行い、実際のデータファイル準備からセルイラストレータオンラインを用いたネットワーク解析までの流れを紹介します。

申込方法：下記フォーマットにて、10 月 21 日（金）17:00 までに e-mail または FAX にてお申込み下さい。

ヒトゲノム解析センター・スーパーコンピューター室

e-mail : hgc-lect@hgc.jp

FAX : 03-5449-5133

「HGC スパコン Web サービス利用法講習会」に参加します。

氏名：

所属：

e-mail：

TEL：

ご要望：

Dear All Users,

On October 25 (Tuesday), we will be holding a workshop titled "Using Web Services", in which we will introduce, in omnibus format, the web services that can be used at the Human Genome Center.

The workshop will be held in Japanese. If you are interested in attending, please contact us at hgc-lect@hgc.jp.

System Administrator
Super Computer System
Human Genome Center