

ヒト遺伝子の統合データベース公開

ヒト完全長 cDNA の配列データに対する包括的なアノテーション（注釈付け）の結果に基づき、ヒト遺伝子の統合データベースである H-Invitational Database（略称 H-InvDB）を完成させた。

H-InvDB は 21,037 個のヒト遺伝子に関するさまざまな情報を蓄えている。その情報はすべて計算機によるバイオインフォマティクス解析と研究者による精査を経たものであり、信頼性が高い。H-InvDB が含む情報は、例えば、遺伝子のゲノム上での位置と構造、タンパク質に翻訳される ORF の位置と翻訳配列、タンパク質としての機能・細胞内局在・立体構造の予測、単塩基置換による遺伝子の多様性、人体の各組織における遺伝子発現パターン、比較ゲノム解析による分子進化の特徴などである。このように広範囲にわたるヒト遺伝子のアノテーション情報を持つデータベースは、世界でも前例がない。

このデータベースは、産総研、バイオ産業情報化コンソーシアム、国立遺伝学研究所を中心とした、世界の 44 カ国の研究者による国際共同研究の成果であり、2004 年 4 月からインターネット*上で公開を開始している。

ヒト完全長 cDNA クローンの生産における日本の貢献率は非常に高く、世界の約 60% を占める。特に経産省主導による NEDO/FLJ プロジェクトでは大量の cDNA クローンが作成されたが、この情報が本プロジェクトの核となった。そして五條堀孝（産総研、遺伝研）、菅野純夫（東大、産総研）、野村信夫（産総研）と筆者らが中心となって本プロジェクトの構想を固め、アメリカ NIH の MGC、ドイツがんセンターの DKFZ、中国上海ゲノムセンターに協力を要請し、世界から合計約 42,000 本のヒト完全長 cDNA のデータを収集した。こうして、2002 年 8 月に東京・臨海副都心において 10 日間の「アノテーション・ジャンボリー会議」を開催し（図 2）、H-InvDB の基盤を築いた。この会議を経て、最新のヒトゲノム配列を用いた再解析や全データの見直し等の作業を実施し、データベースの完成に至った。

H-InvDB が、ヒトのトランスクリプトーム研究の基盤として生命の仕組みを理解するのに役立つことはもちろんのこと、癌などの病因解明や創薬のための基礎情報の提供など将来にわたり幅広く活用され役立つこと期待している。

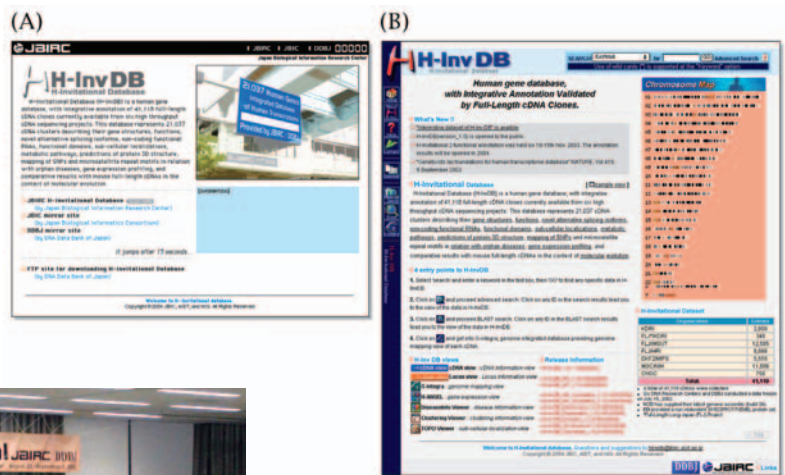


図 1 H-Invitational Database の (A) 公式ページと (B) データベース本体 (右)



図 2 H-Invitational ジャンボリーの模様 (左下)



いまにし ただし
今西 規
imanishi@jbirc.aist.go.jp
生物情報解析研究センター

関連情報

- H-Invitational Database, <http://www.h-invitational.jp/>
- T. Imanishi, T. Itoh, Y. Suzuki, C.O'Donovan, S. Fukuchi, K.O.Koyanagi, R.A.Barrero, T. Tamura, Y. Yamaguchi-Kabata, M. Tanino, et al: PLoS Biology Vol. 2, No. 6, 856-875 (2004).
- D. Cyranoski: Nature, Vol. 419, 3-4 (2002).
- *<http://www.h-invitational.jp> (または hin.jp) にて利用できる (図1)。