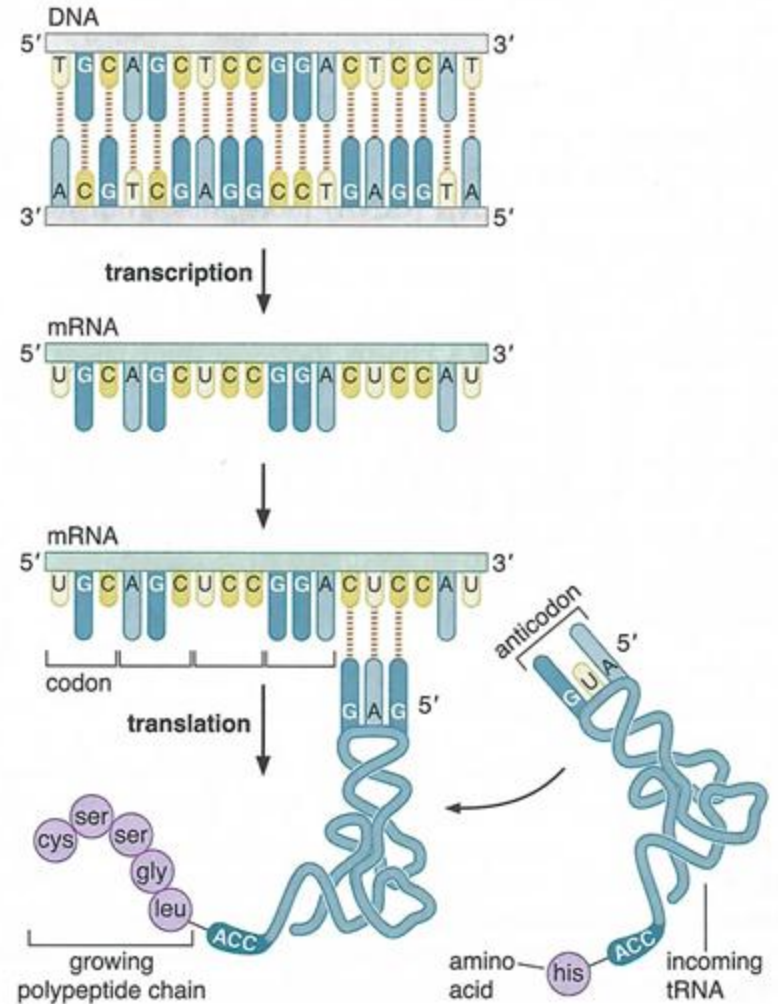


# Central Dogma

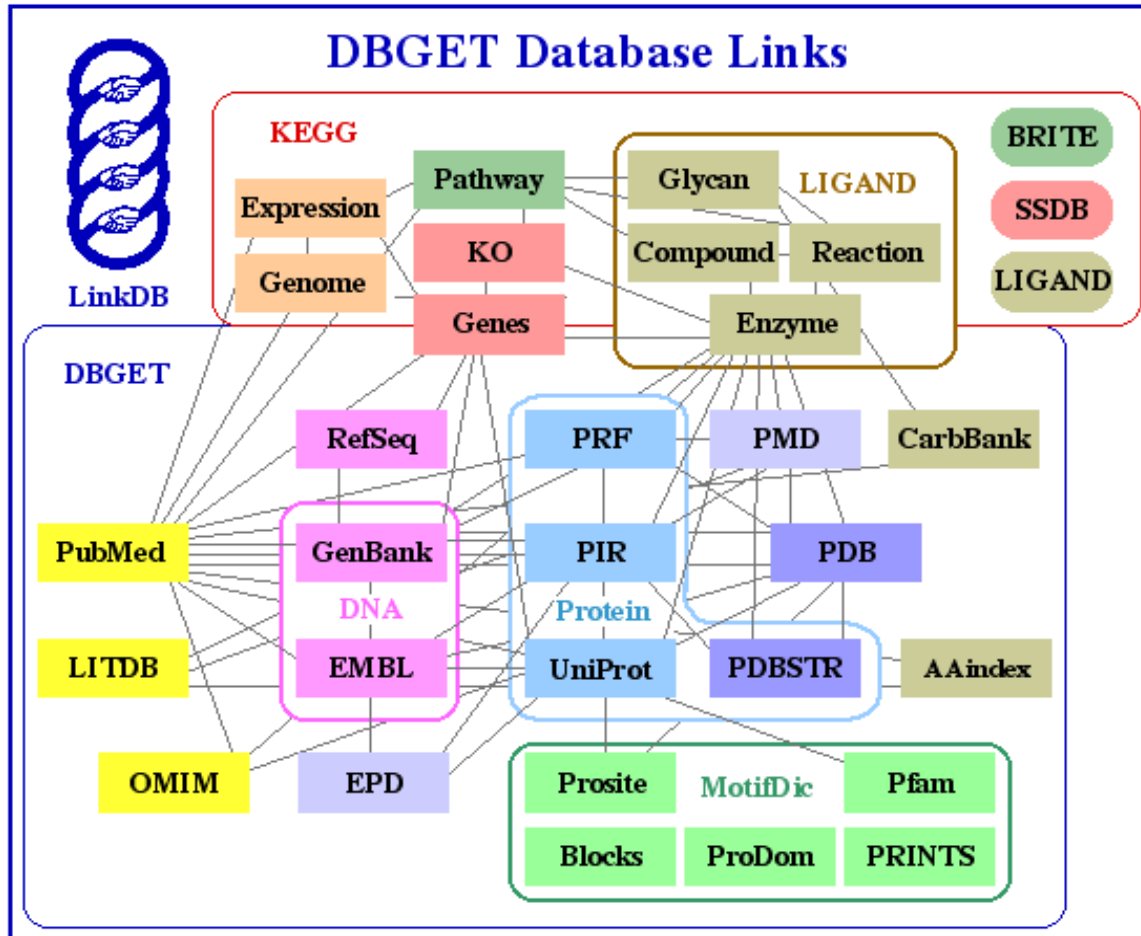
- DNA → RNA → Protein

転写

翻訳



# GenomeNet



**DB Entry:**  
Gene or Protein  
Research Group

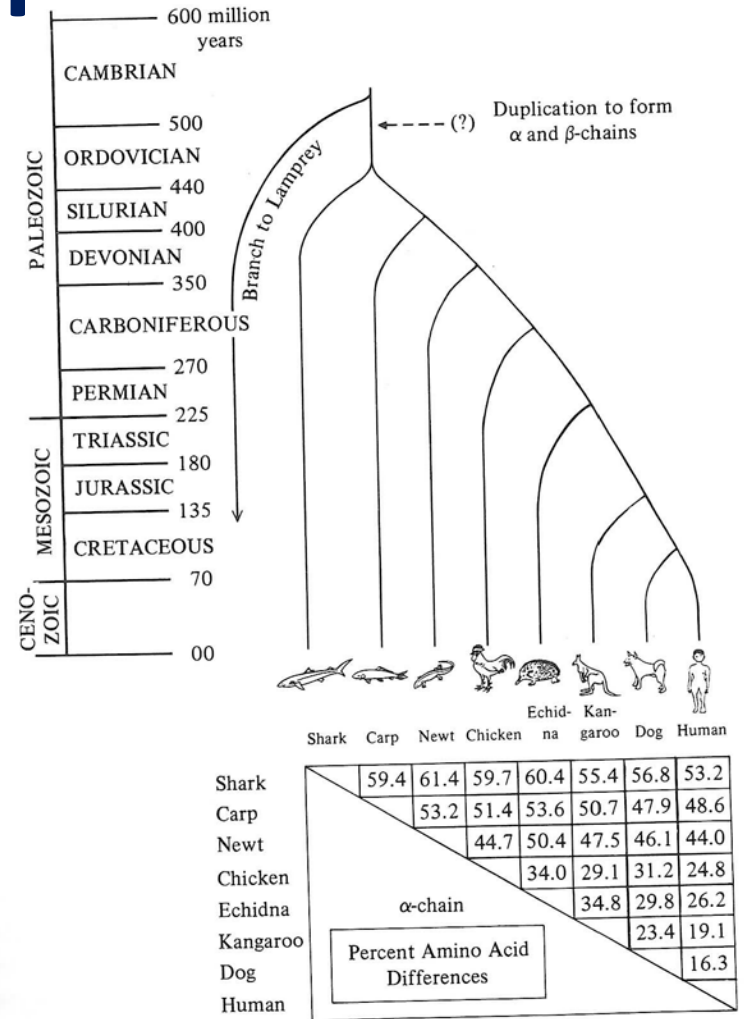
Click on the following: to invoke the following:

Database name	Basic DBGET search
DBGET	Advanced DBGET search
KEGG	KEGG table of contents
LinkDB	LinkDB search



# Molecular Evolution

Fig. 4.2. Percentage amino acid differences when the  $\alpha$  hemoglobin chains are compared among eight vertebrates together with their phylogenetic relationship and the times of divergence.



駒井 卓 (1886—1972)  
 地球の歴史は地層に  
 生物の歴史は分子に

Kimura (1983)

Volume 5

Supplement I

1973

# ATLAS of PROTEIN SEQUENCE and STRUCTURE

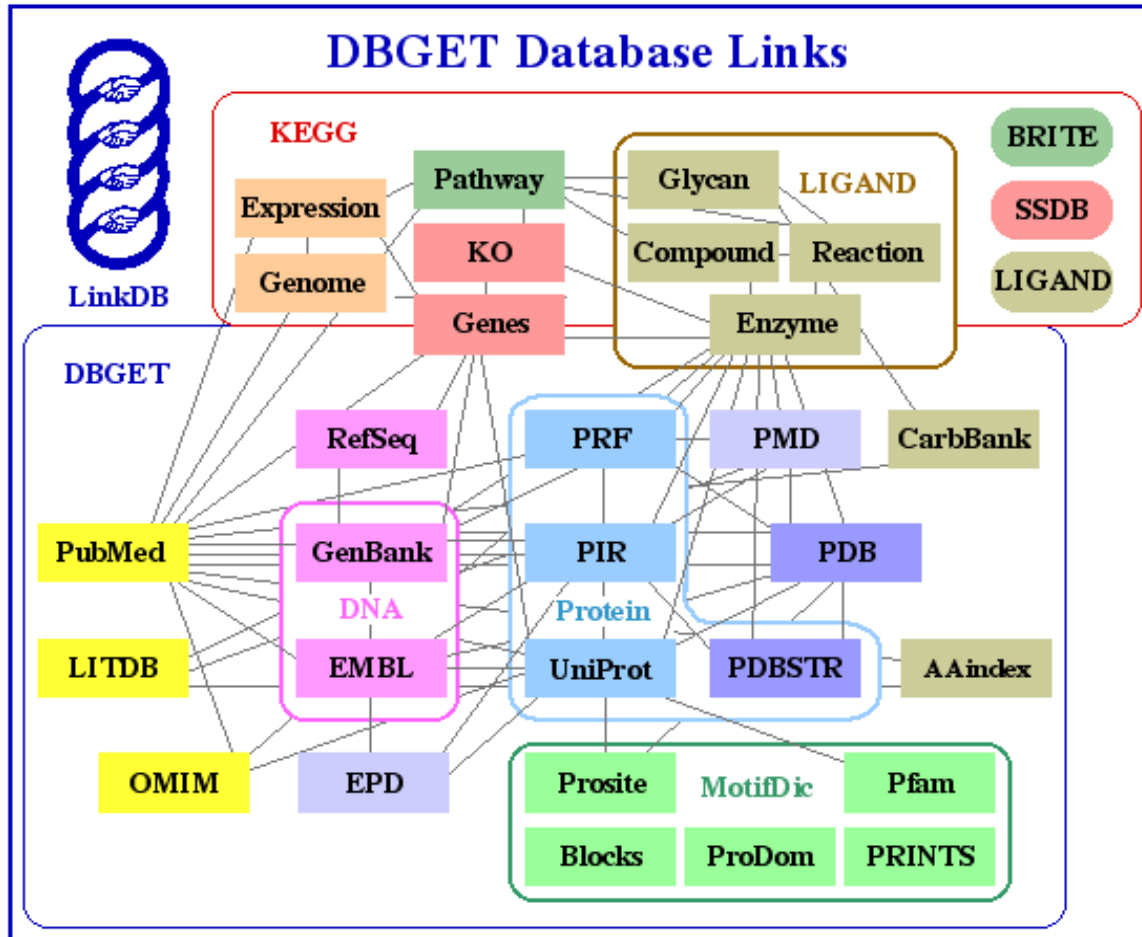
HIS ARG  
GLU ALA TRP LEU PRO LYS  
PRO DAY HDF FER LEU CYS PHE SER GLE SER  
ASN ARG GLU GLU THR GLN LYS BAR KER LEU  
SER LEU LEU SHI ULE GLN SER TRP LEI  
ARG SER VAL PHE ALA ASN SER  
ASP HUN TYR ASP LEU LEU  
LEU MET GLY ARG LEU GIN  
E LYS GLN THR TYR  
ALA LEU LEU  
TYR ASP  
ASN TYR  
VAL GLU THR  
GLY SER CYS GLY  
LEU LYS GLU ASN PAR  
ALA ASN ILE PRO GLN  
LYS ALA PRO ALA ALA LYS  
THR ASX GLY VAL ASX PRO  
LEU PHE ALA LEU VAL ARG ASP  
GLY THR VAL VAL AMC LAU GHL  
ALA GLN LYS ALA VAL THR ASN PRO GLU  
LEU LEU LYS THR TYR LYS ALA ALA VAL  
LEU SER ARG TYR LEU VAL AN

Margaret O. Dayhoff



NATIONAL BIOMEDICAL RESEARCH FOUNDATION  
GEORGETOWN UNIVERSITY MEDICAL CENTER  
WASHINGTON, D. C. 20007

# GenomeNet



**Submission  
to DB  
(Accession  
No.)**

**DB Entry:  
Gene or Protein  
Research Group**

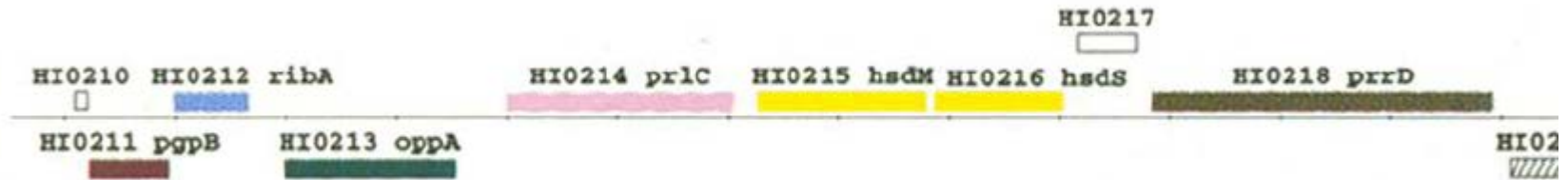
Click on the following: to invoke the following:

Database name	Basic DBGET search
DBGET	Advanced DBGET search
KEGG	KEGG table of contents
LinkDB	LinkDB search

# The Era of Genomics

## *Haemophilus influenzae* (1995)

1,830,138 bp, 1657 proteins



GenomeNet KEGG (March 2013)

Eukaryotes: 180 Bacteria: 2149 Archaea: 149

# Human Genome

articles

## Finishing the euchromatic sequence of the human genome

International Human Genome Sequencing Consortium\*

*\* A list of authors and their affiliations appears in the Supplementary Information*

---

The sequence of the human genome encodes the genetic instructions for human physiology, as well as rich information about human evolution. In 2001, the International Human Genome Sequencing Consortium reported a draft sequence of the euchromatic portion of the human genome. Since then, the international collaboration has worked to convert this draft into a genome sequence with high accuracy and nearly complete coverage. Here, we report the result of this finishing process. The current genome sequence (Build 35) contains 2.85 billion nucleotides interrupted by only 341 gaps. It covers  $\sim 99\%$  of the euchromatic genome and is accurate to an error rate of  $\sim 1$  event per 100,000 bases. Many of the remaining euchromatic gaps are associated with segmental duplications and will require focused work with new methods. The near-complete sequence, the first for a vertebrate, greatly improves the precision of biological analyses of the human genome including studies of gene number, birth and death. Notably, the human genome seems to encode only 20,000–25,000 protein-coding genes. The genome sequence reported here should serve as a firm foundation for biomedical research in the decades ahead.

Draft (2001)  $\rightarrow$  Nature 431, 931-945 (2004)



# Individual Human Genomes

1. Craig Venter
2. Jim Watson
3. 4. 5. ....

Whole-genome sequencing and comprehensive variant analysis of a Japanese individual using massively parallel sequencing

Akihiro Fujimoto<sup>1,2</sup>, Hidewaki Nakagawa<sup>1</sup>, Naoya Hosono<sup>1</sup>, Kaoru Nakano<sup>1</sup>, Tetsuo Abe<sup>1</sup>, Keith A Boroevich<sup>1</sup>, Masao Nagasaki<sup>3</sup>, Rui Yamaguchi<sup>3</sup>, Tetsuo Shibuya<sup>3</sup>, Michiaki Kubo<sup>1</sup>, Satoru Miyano<sup>2,3</sup>, Yusuke Nakamura<sup>1,3</sup> & Tatsuhiko Tsunoda<sup>1,2</sup>

**Nat Genet. 42, 931-936 (2010)**

# 1000 Genomes

A Deep Catalog of Human Genetic Variation



[Home](#) [About](#) [Data](#) [Analysis](#) [Participants](#) [Contact](#) [Browser](#) [Wiki](#) [FTP search](#)

 Search

## LATEST ANNOUNCEMENTS

WEDNESDAY OCTOBER 31, 2012

### An integrated map of genetic variation from 1092 human genomes

The Phase 1 publication, [An Integrated map of genetic variation from 1092 human genomes](#) is now available from [Nature](#) and can be downloaded directly from the [ftp site](#). The paper is distributed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported licence. Please share our paper appropriately.

All the data files associated with this paper can be found in our [phase1 analysis results directory](#).

## Recent project announcements

WEDNESDAY MARCH 06, 2013

### New sequence index released

A new sequence index is now available on the FTP site

[http://ftp.1000genomes.ebi.ac.uk/vol1/ftp/sequence\\_indices/20130305.sequence.index](http://ftp.1000genomes.ebi.ac.uk/vol1/ftp/sequence_indices/20130305.sequence.index)

## NAVIGATION

- [Frequently Asked Questions](#)

## LINKS



[All Project Announcements](#)



[Sample and Project Information](#)



[Media Archive](#)



[Download the 1000](#)

23andMe can help you manage risk and make informed decisions...



Ancestry

Connect to your past.



Health

Learn for the present.



Research

Participate for the future.

Learn valuable health & ancestry information.

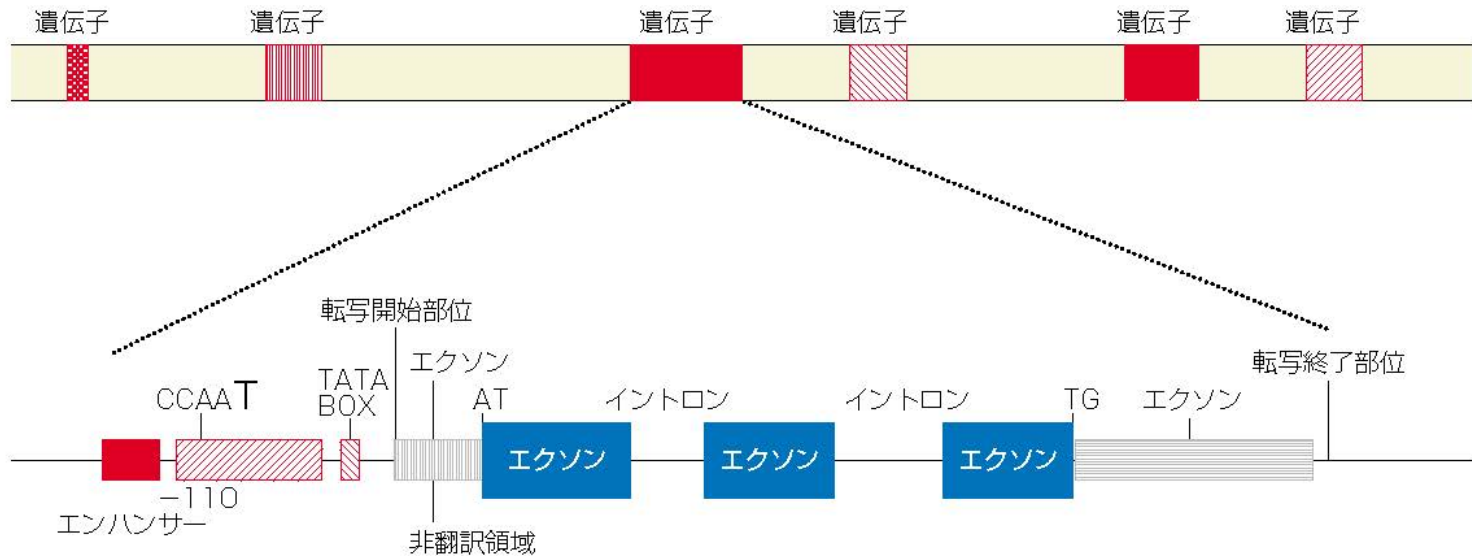


For only \$99



**2015 15min \$100**

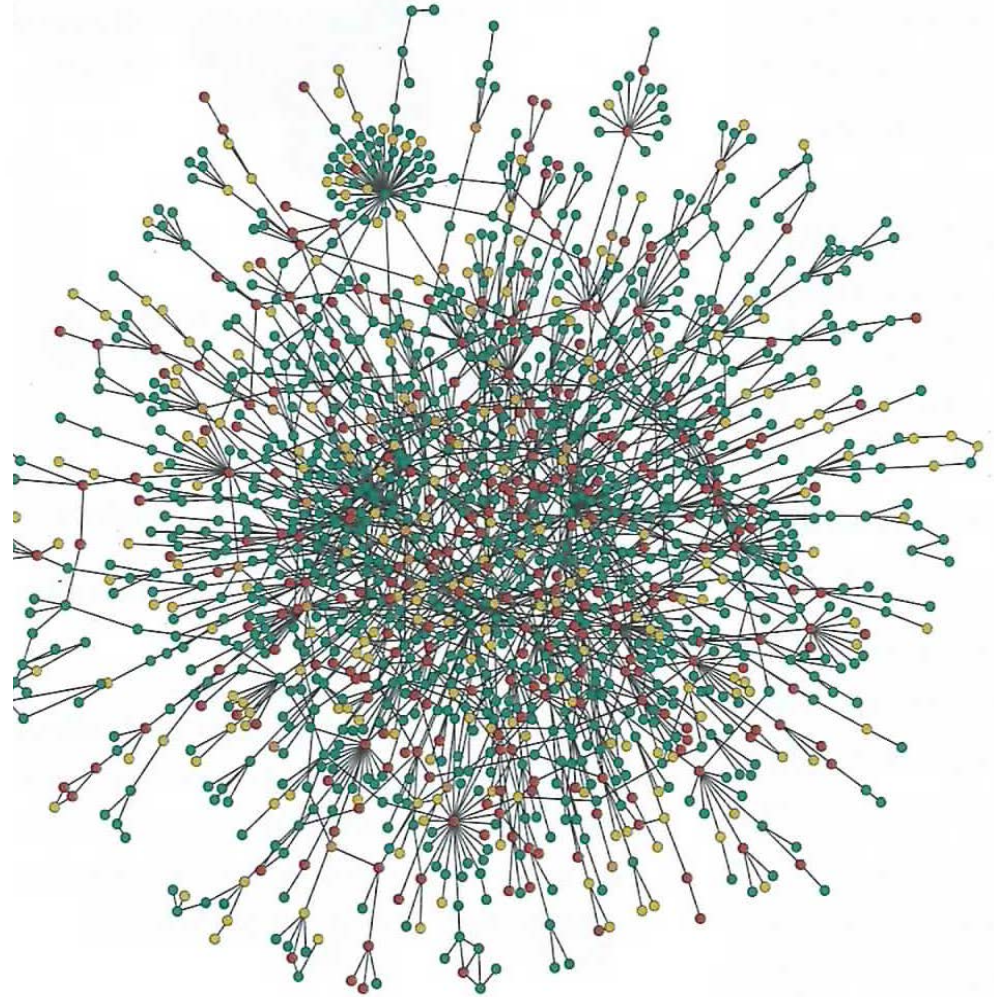
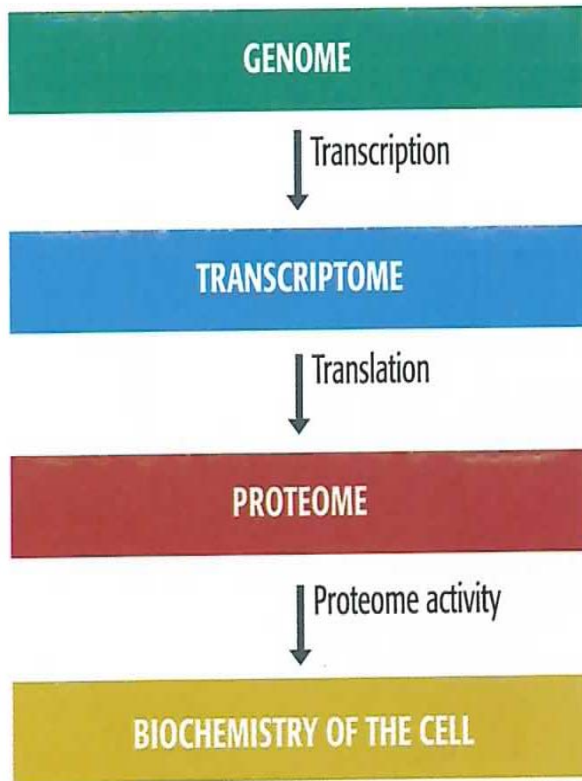
# Annotations to Genomic Data



- Coding Regions
- Transcription Initiation
- Epigenetics
- SNPs

} Regulation

# Gemonics and Proteomics



Nature (2001)

[\[English\]](#)

H-InvDBとは？

[プロジェクト](#)[論文](#)[学会発表](#)[更新情報](#)[Webサービス](#)[アノテーション・トピックス](#)[遺伝子ファミリー・グループ](#)[ニュース](#)[統計情報](#)[メールマガジン](#)[メンテナンス情報](#)[よくある質問](#)

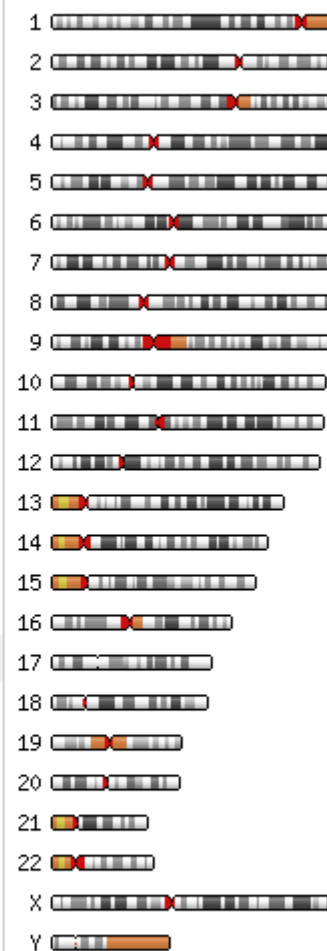
## ヒト遺伝子アノテーション統合データベース

H-InvDBはヒトの遺伝子と転写産物を対象とした統合データベースです。ヒトのすべての転写産物の配列をあらゆる手法で解析することにより、[ヒト遺伝子の構造](#)、[選択的スプライシングバリエーション](#)、[機能性RNA](#)、[タンパク質としての機能](#)、[機能ドメイン](#)、[細胞内局在](#)、[代謝経路](#)、[立体構造](#)、[疾病との関連](#)、[遺伝子多型 \(SNP、マイクロサテライト等\)](#)、[遺伝子発現プロファイル](#)、[分子進化学的特徴](#)、[タンパク質間相互作用 \(PPI\)](#)、[遺伝子ファミリー](#)などの精査されたアノテーション (注釈付) 情報を提供しています。H-InvDBは、H-Invitationalプロジェクトで確立したヒト完全長cDNA配列のアノテーション技術を基礎として、ゲノム情報統合プロジェクト (2005-2008)、経済産業省統合データベースプロジェクト (2008-2011)、経済産業省ライフサイエンスデータベースプロジェクト (2008-2011) の主要データベースとして、構築・更新されています。

[\[more\]](#)

## ニュース




- 2012-12-21 H-InvDB\_8.0 データ修正
- 2012-12-21 H-InvDBに関する論文出版 (NAR database issue 2013)
- 2012-12-21 タンパク質複合体データベース「略称PCDq」をリリース
- 2012-12-21 ヒトタンパク質データベース「略称H-EPD」をリリース
- 2012-07-31 [「平成24年度経済産業省ライフサイエンス関連データベースに関するアンケート\(2012\)」の募集を開始しました。](#)
- 2012-04-20 ヒト全遺伝子アノテーションデータベース「H-InvDB 8.0」をリリース
- 2010-09-10 Erratum H-InvDB\_7.5 Evolaアノテーション情報
- 2010-09-10 ヒト全遺伝子アノテーションデータベース「H-InvDB 7.5」をリリース
- 2010-02-16 ヒト全遺伝子アノテーションデータベース「H-InvDB 7.0」をリリース
- 2010-01-14 「H-InvDB遺伝子リスト特徴抽出ツール (略称HEAT)」更新









# H-InvDB

## ■ Databases, viewers and tools in H-InvDB







### ▶ データベース

-  Locus view [Help] [Sample view]
-  Transcript view [Help] [Sample view]
-  Protein view [Help] [Sample view]

### ▶ サブデータベース

-  G-integra [Top] [Help] [Sample view]
-  H-ANGEL [Top] [Help] [Sample view]
-  DiseaseInfo Viewer [Help] [Sample view]
-  Evola [Top] [Help] [Sample view]
-  PPI view [Top] [Help] [Sample view]
-  Gene Family/Group [Top] [Help] [Sample view]

### ▶ サテライトデータベースツール

-  H-DBAS [Top] [Help] [Sample view]
-  VarySysDB [Top] [Help] [Sample view]
-  G-compass [Top] [Help]
-  LEGENDA [Top]
-  TACT [Top] [Help] [Sample view]
-  HEAT [Top] [Help]

## ■ Information and documents about H-InvDB

### ▶ 更新情報

- H-InvDB release information
- H-InvDB release history

### ▶ 統計情報

- H-InvDB statistics

### ▶ アノテーション方法

- ヒト転写物のクラスター [Help]
- 選択的スプライシング変異体 [Help]
- タンパク質コード遺伝子 [Help]
- 非タンパク質コード遺伝子 [Help]
- 分子進化解析 [Help]

### ▶ ドキュメント

- H-InvDBプロジェクトの概要 [Help]
- 関連文献情報 [Help]
- H-InvDB学会発表 [Help]
- ニュース [Help]
- メンテナンス情報 [Help]
- よくある質問(FAQs) [Help]

# Database から Knowledgebase

Big Data

Machine Annotation

Fixed Data  Reorganization of Knowledge

Error Propagation in the current DB system