

# ヒトゲノム解読から10年、黒船来航から160年 ～日本は夢物語、米国では明日の問題～

## NPO健康医療開発機構GDHD

宮野 悟

NPO健康医療開発機構 理事・事務局長

東京大学医科学研究所ヒトゲノム解析センター

miyano@ims.u-Tokyo.ac.jp

「世界に一つだけの花」  
私たちは一人一人異なる  
ゲノムを持っている

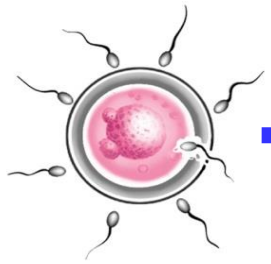


「がんは日本の国民病」日本人  
の半分以上が罹り、3分の1が亡く  
なっている



# 「私のDNA」と「私」

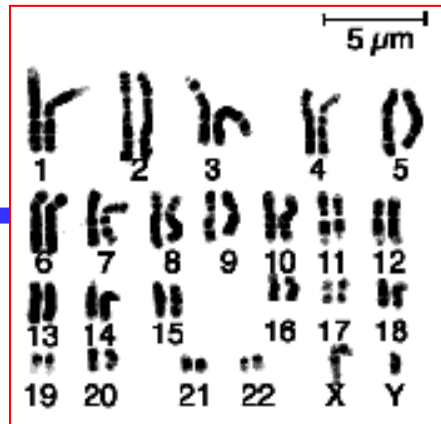
母のDNAと卵細胞



父のDNA

1954年

23組の染色体  
30億×2文字の  
ATCGのDNA情報



環境因子

生命のメカニズム

誕生、成長  
結婚、子育て  
病気  
ボケ、がん

死

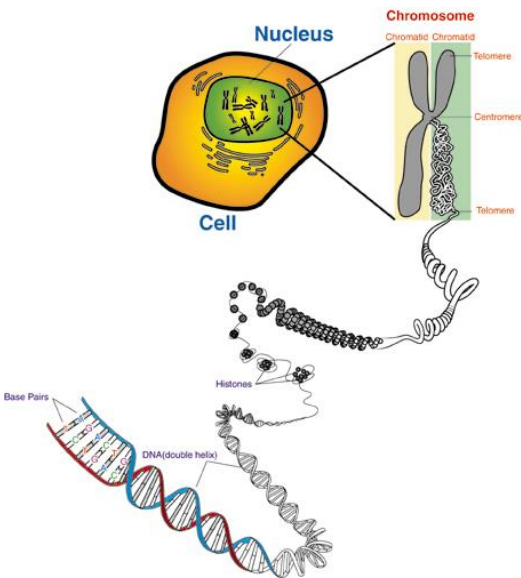
2013年



# 皆さんの人生(誕生、成長、結婚、子育て、そしてやがてやってくるボケ、がん、死)とは切ってもきれない「退屈な話」が分子生物学のセントラルドグマです

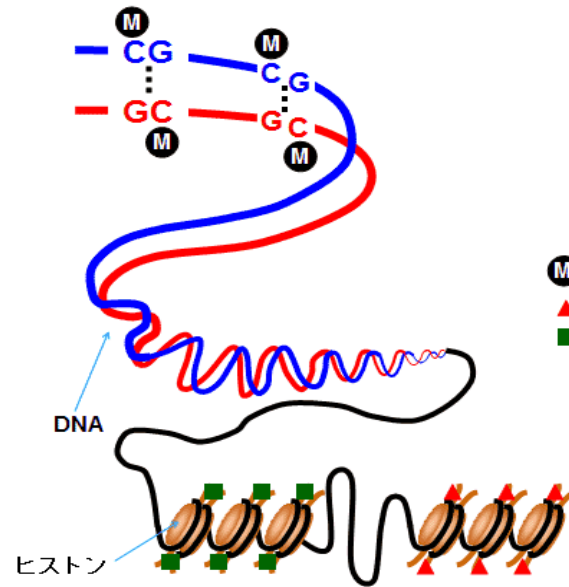
DNA

ゲノム



DNAやヒストンの修飾状態

エピゲノム

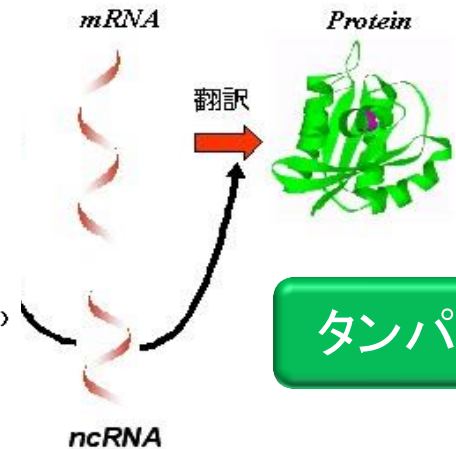


遺伝子発現データ

mRNAやnoncoding RNA

トランスクリプトーム

M DNAにつく目印 (DNAメチル化)  
▲ ヒストンにつく目印 (ヒストン修飾)



タンパク質

タンパク質に翻訳されないDNAからの転写産物

Image source:  
[www.biotec.or.th/Genome/whatGenome.html](http://www.biotec.or.th/Genome/whatGenome.html)

分子生物学のセントラルドグマ

TOP STORIES IN U.S.

1 of 12



Massive Online Courses Woo Professors



New Tack in Preventing Hospital Infections

2 of 12



How Three Small Firms Face Health Law

3 of 12

Sus  
to Ob

U.S. NEWS | Updated May 14, 2013, 9:25 a.m. ET

## Angelina Jolie Says She Had Double Mastectomy

Article

Video

Comments (25)

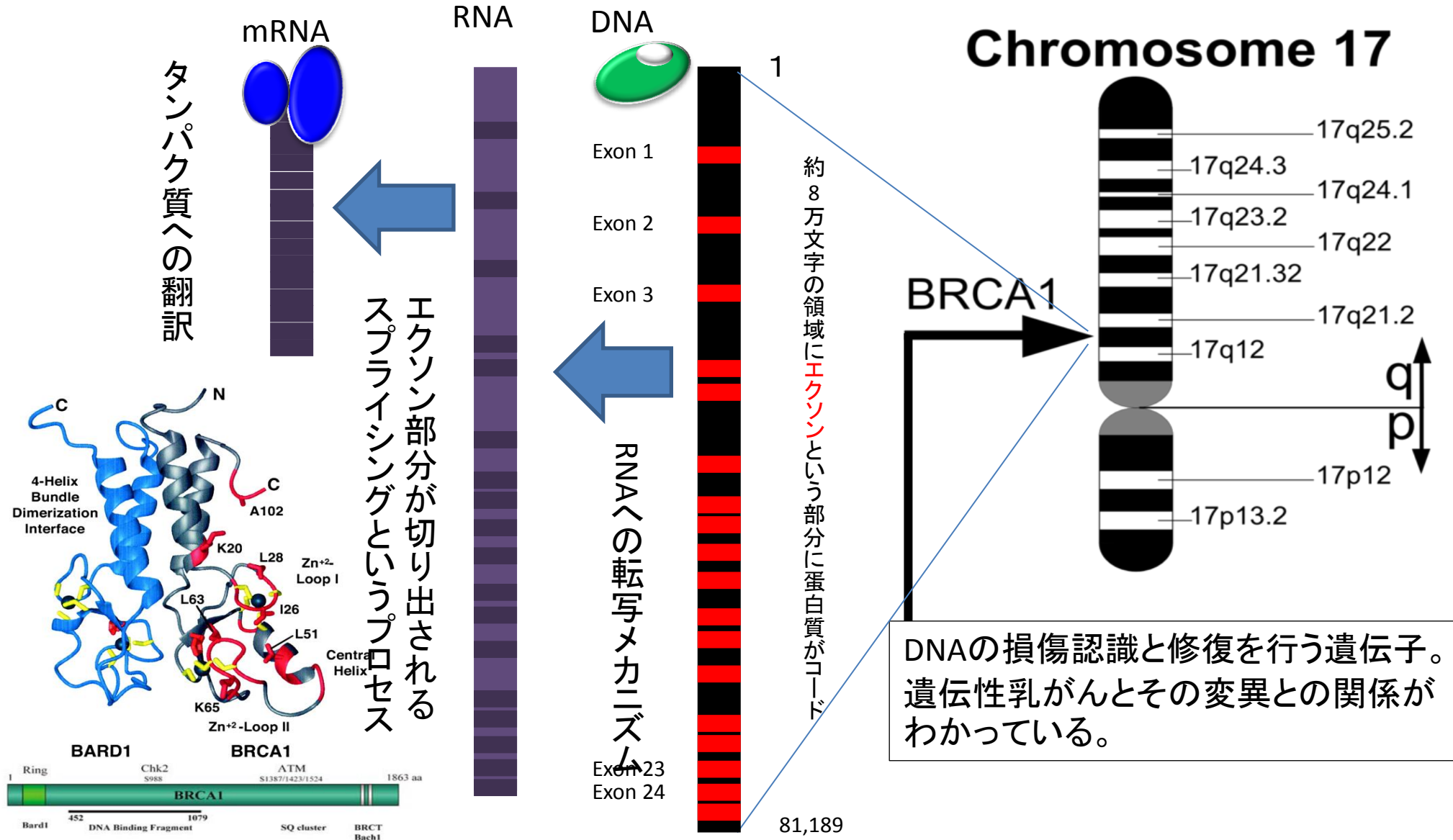
Associated Press



<http://www.youtube.com/watch?v=1qW3zwqp-kY>

# ヒト17番染色体

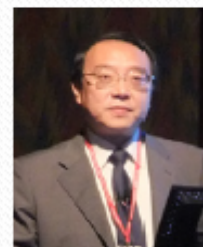
約8100万文字のDNA情報



# 乳がんの分子サブタイプ分類と 個別化抗がん剤治療の開発

三木 義男<sup>1</sup> (研究代表者)  
松浦 正明<sup>2</sup>

1 東京医科歯科大学難治疾患研究所分子遺伝分野  
2 (公財)がん研究会がん研究所がんゲノム研究部



SCIENCE • VOL. 266 • 7 OCTOBER 1994

RESEARCH ARTICLES

## A Strong Candidate for the Breast and Ovarian Cancer Susceptibility Gene **BRCA1**

Yoshio Miki, Jeff Swensen, Donna Shattuck-Eidens, P. Andrew Futreal, Keith Harshman, Sean Tavtigian, Qingyun Liu, Charles Cochran, L. Michelle Bennett, Wei Ding, Russell Bell, Judith Rosenthal, Charles Hussey, Thanh Tran, Melody McClure, Cheryl Frye, Tom Hattier, Robert Phelps, Astrid Haugen-Strano, Harold Katcher, Kazuko Yakumo, Zahra Gholami, Daniel Shaffer, Steven Stone, Steven Bayer, Christian Wray, Robert Bogden, Priya Dayananth, John Ward, Patricia Tonin, Steven Narod, Pam K. Bristow, Frank H. Norris, Leah Helvering, Paul Morrison, Paul Rosteck, Mei Lai, J. Carl Barrett, Cathryn Lewis, Susan Neuhausen, Lisa Cannon-Albright, David Goldgar, Roger Wiseman, Alexander Kamb, Mark H. Skolnick\*

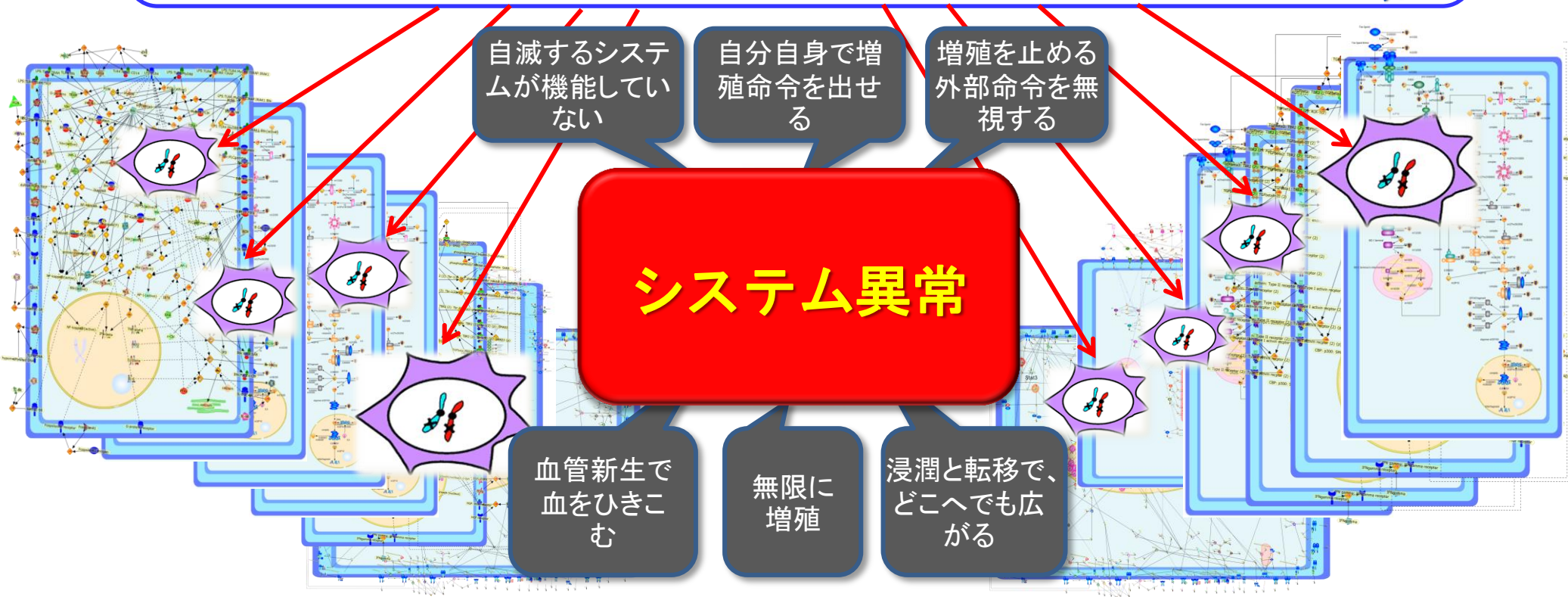
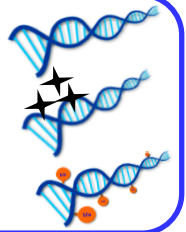
A strong candidate for the 17q-linked *BRCA1* gene, which influences susceptibility to breast and ovarian cancer, has been identified by positional cloning methods. Probable predisposing mutations have been detected in five of eight kindreds presumed to segregate *BRCA1* susceptibility alleles. The mutations include an 11-base pair deletion, a 1-base pair insertion, a stop codon, a missense substitution, and an inferred regulatory mutation. The *BRCA1* gene is expressed in numerous tissues, including breast and ovary, and encodes a predicted protein of 1863 amino acids. This protein contains a zinc finger domain in its amino-terminal region, but is otherwise unrelated to previously described proteins. Identification of *BRCA1* should facilitate early diagnosis of breast and ovarian cancer susceptibility in some individuals as well as a better understanding of breast cancer biology.

「システムがん」班員の  
三木義男(東京医  
科歯科大学・教授)が  
見つけた遺伝子

# 私はなぜがんになるのか？



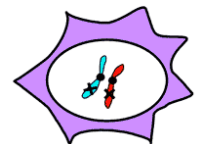
- 個人個人で異なっている遺伝的要因 (DNA)
- 腫瘍細胞に蓄積した遺伝子変異 (がんゲノム)
- 環境要因などによるゲノム修飾の変異 (エピゲノム)



正常細胞



「私」のゲノムの情報が不可欠



がん細胞

# My DNAを「シーケンス」

生物のDNA情報を読み取る装置は一般にシーケンサーとよばれ、A, T, C, Gの文字で綴られるDNA情報(ヒトの場合30億文字の情報)をコンピュータで読めるように取り出すことを「シーケンス」とよんでいます。



# だれもが自分のDNA情報を利用できる時代が到来

現在2013年

27時間、40万円  
装置：8000万円  
リード長：100-150



Illumina HiSeq2500

限界 シークエンス量  
に線形比例して高価  
な試薬が必要

2013年～2014年

6時間、10万円  
装置：1500万円  
リード長：400b



Ion Proton

2014年

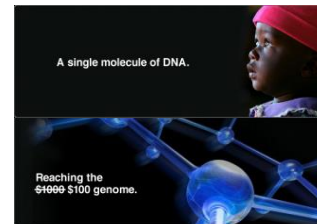
1時間、10万円～  
装置：300-6000万円  
リード長：50Kb-100Kb



Oxford Nanopore Technologies  
2000ナノポア集積

2014年～

1時間以内、1万円  
装置：10万円  
リード長：1Mb



Genia Technologies  
100万ナノポア集積  
COMSチップ→数百円

2015年～

超安価・高速・  
高精度

IBMなどを初め  
とする様々な技  
術開発と応用へ  
の展開  
ゲノム及びそれ  
に由来する情報  
の臨床翻訳・解  
釈が最も重要な  
課題

★ **シリコンシーケンサーの登場**  
シリコンチップのコストはとても安い

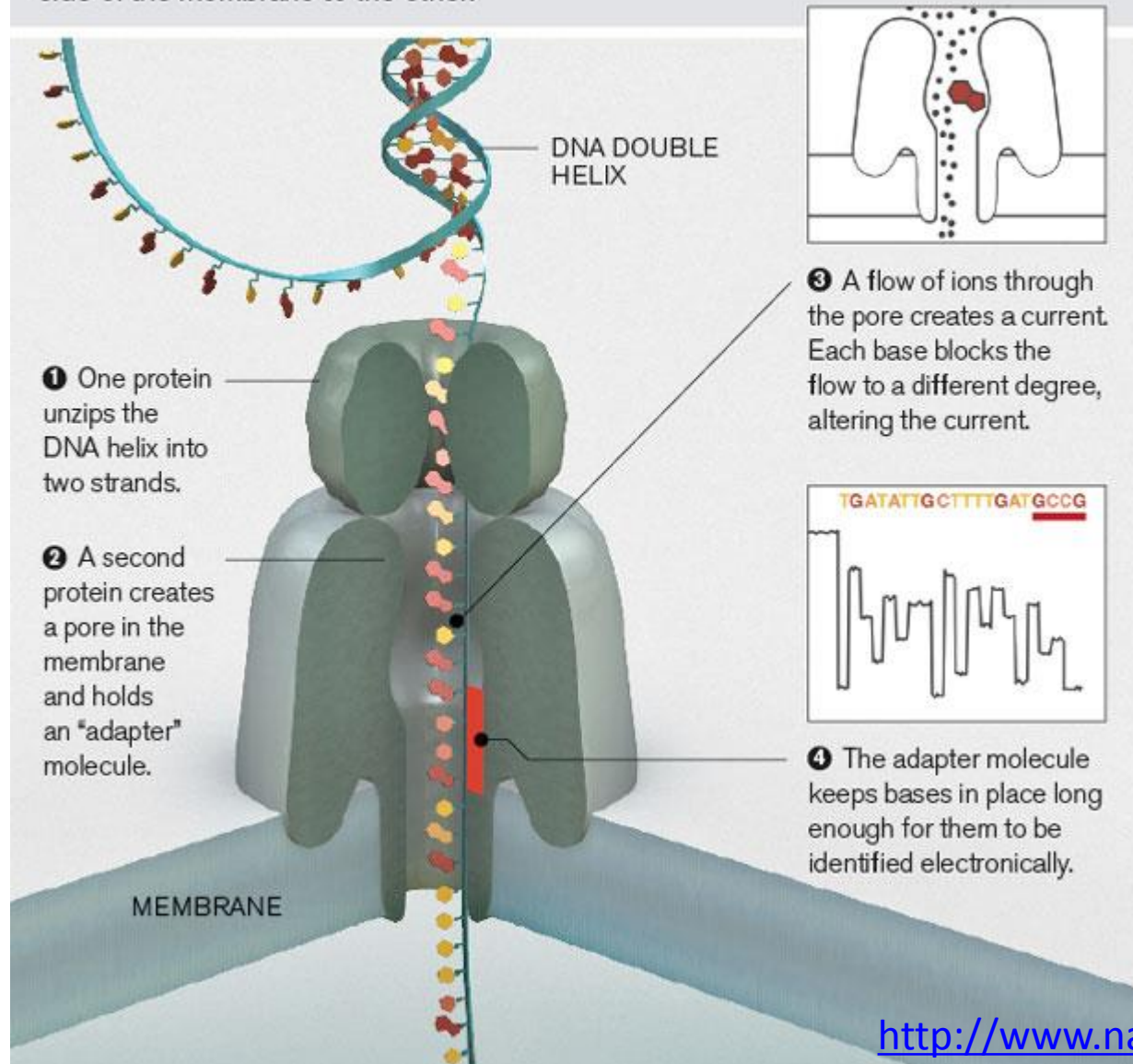
さらに、トランスクリプトーム、  
エピゲノムもこの装置で解析

しかし、DNAビッグデータを解読して  
活用できる人材・計算資源が不足



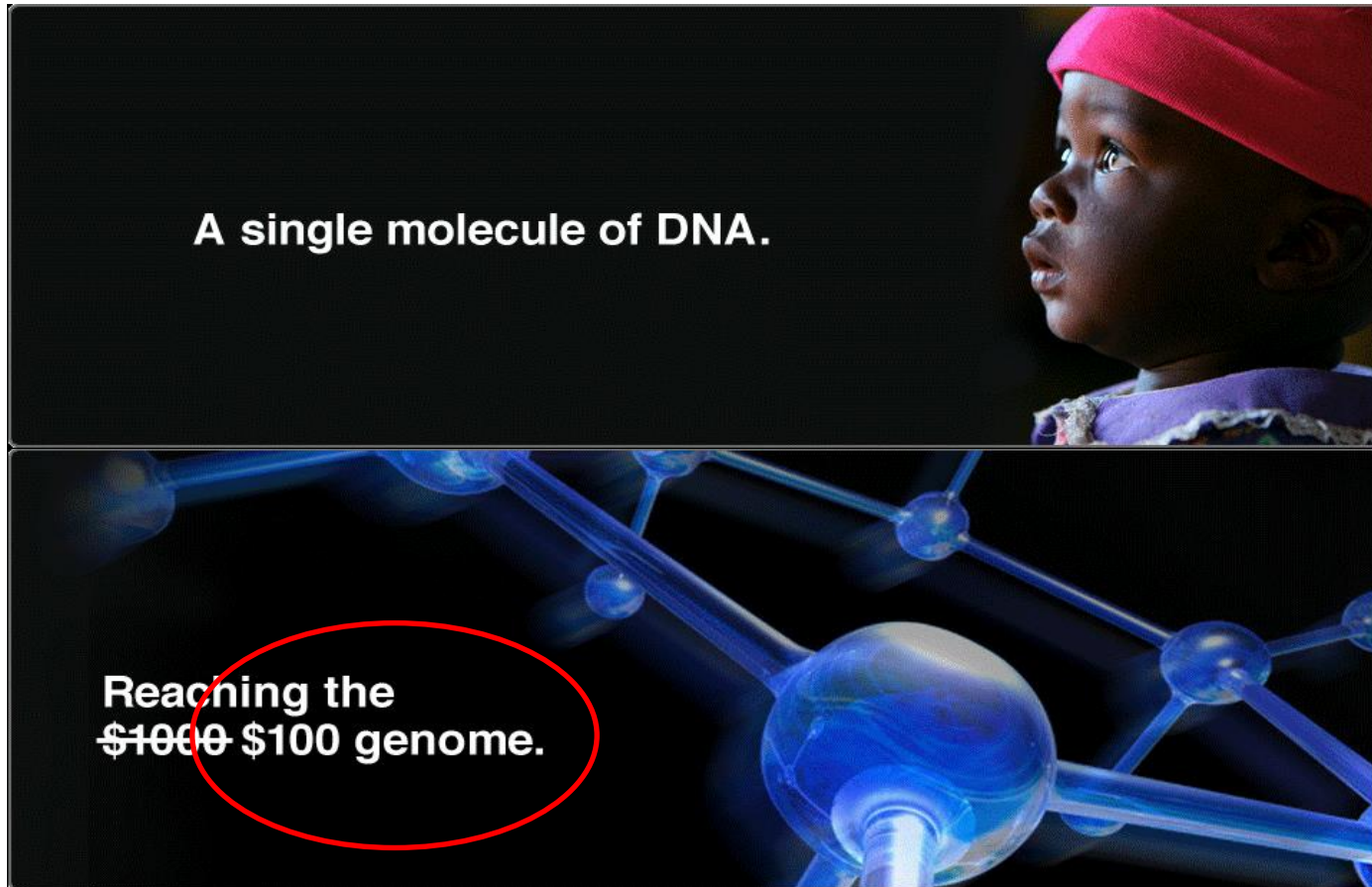
# Oxford Nanopore Technologies

DNA can be sequenced by threading it through a microscopic pore in a membrane. Bases are identified by the way they affect ions flowing through the pore from one side of the membrane to the other.



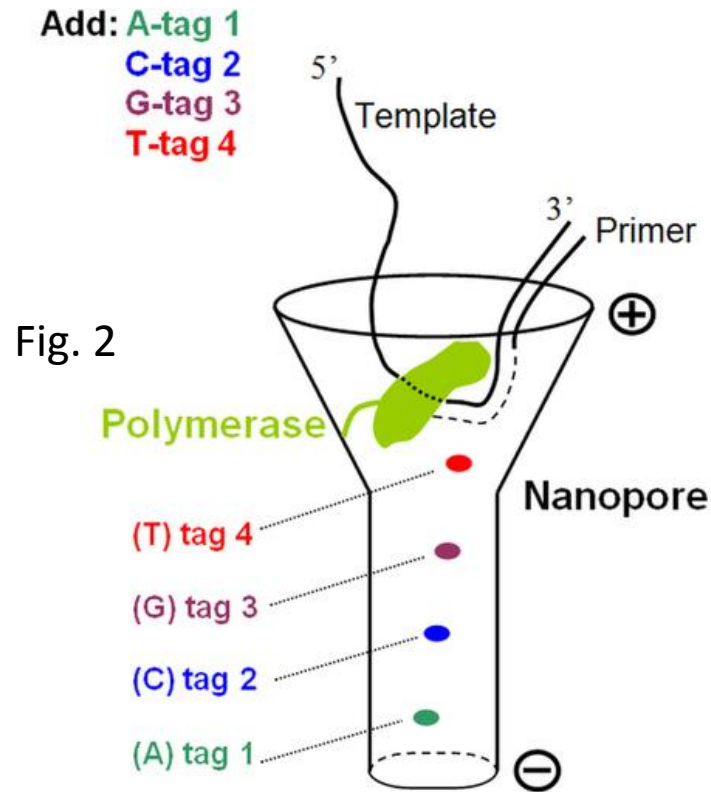
# Genia Technologies, Inc.

2013年～ヒトゲノムは100ドル以下、1時間以内  
“Last-Generation Sequencer”



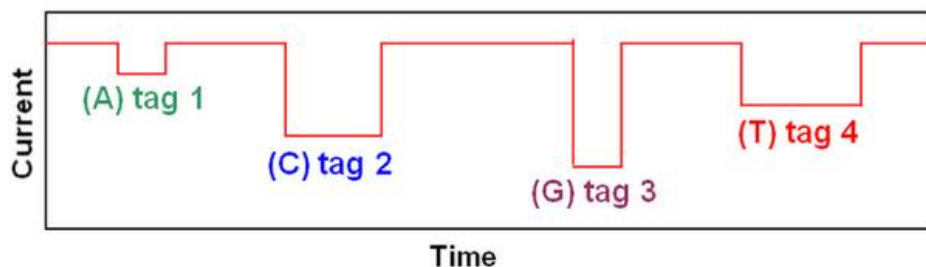
<http://www.geniachip.com/>

# DNA sequencing by a nanopore with phosphate-tagged nucleotides



## General principle of single molecule electronic DNA sequencing by synthesis using PEG-labeled nucleotides and nanopore detection

The single molecule electronic Nano-SBS system, which is shown schematically in [Fig. 2](#), depicts the DNA polymerase bound in close proximity to the nanopore entrance. A template to be sequenced is added along with the primer. To this template-primer complex, four differently tagged nucleotides are added to the bulk aqueous phase. After polymerase catalyzed incorporation of the correct nucleotide, the tag-attached polyphosphate will be released and pass through the nanopore to generate a unique ionic current blockade signal, thereby identifying the added base electronically because the tags have distinct chemical structures.



Kumar S, Tao C, Chien M, Hellner B, Balijepalli A, Robertson JW, Li Z, Russo JJ, Reiner JE, Kasianowicz JJ, Ju J. PEG-labeled nucleotides and nanopore detection for single molecule DNA sequencing by synthesis. *Scientific Reports* 2, Article number: 684, 2012. doi:10.1038/srep00684  
<http://www.nature.com/srep/2012/120921/srep00684/full/srep00684.html>

# 「高速に大量のシーケンスデータはでてくるが・・・この解析がボトルネック」

- パーソナルゲノム時代が到来し、米国ではNIH・民間・医師会が共同し、個別化ゲノム医療の実施体制が急速に作られている。
- スパコン、大規模ストレージなどのリソースとデータ解析の人材が必要

2013年1月10日

米国National Institute of Health所長  
Francis Collinsは、“NIH & Big Data”対応に  
Associate Director for Data Scienceを設置し、人  
材をリクルートすることを発表

“Tackling the Big Data Program”

- Centers of Excellence for Biomedical Big Data-

13



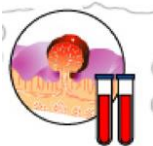
米国NIH所長  
F. Collins

「私のがん」の黒幕たちを暴き出すために、なぜスパコンが必要なのか？

# スーパーコンピュータで シーケンスのアセンブリとデータ解析



## メディカルインフォマ ティクスの達人たち



検体

Ion Proton



半導体シーケンサー

6時間

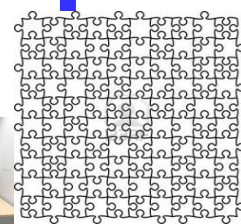
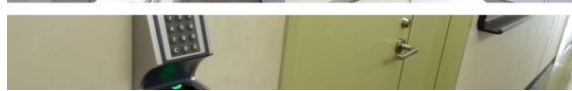
正常組織 900億文字  
がん組織 1200億文字  
のデータ

Human Genome Center  
Institute of Medical Science, University of Tokyo

スパコンシステム



1日~



親から受け  
継いだゲノ  
ム

がんを発症  
する鍵遺伝  
子の変化

環境因子に  
よる変化を  
受けたゲノ  
ム

黒幕  
X

黒幕  
Z

\*\*\* 検査結果 時系列画面プリント \*\*\*

プリント時間： 2012. 4. 9 13:23:35 P. 1

患者番号： [REDACTED] 氏名： [REDACTED]

採取日	2011. 10. 19	2012. 01. 12	2012. 01. 18	2012. 04. 09
採取時間	00:00	00:00	00:00	00:00
総蛋白	6.9~8.4	7.1	7.0	7.8
アルブミン	3.9~5.2	4.1	3.8 L	4.4
CK	62~287	82	***	100
AST (GOT)	13~33	34 H	64 H	45 H
ALT (GPT)	8~42	39	72 H	49 H
LD	119~229	163	173	217
γGT	9~109	101	167 H	127 H
尿素窒素	8~21	13	15	13
クレアチニン	0.6~1.1	0.9	0.8	0.9
eGFR	項目説明参照	71.2	76.6	64.3
尿酸	2.5~7.0	5.6	5.7	6.2
空腹時血糖	70~110(空腹時)	101	96	112
食事日時				
食事内容				
採血日時				
HbA1c (NGSP)	4.6~6.2	***	***	5.7
HbA1c (JDS)	4.3~5.8	5.5	5.5	5.3
中性脂肪	30~150	128	145	144
総コレステロール	122~240	191	200	226
HDLコレステロール	35~70	71 H	77 H	87 H
LDLコレステロール(直接法)	140未満	84	85	99
Na	139~146	142	142	141
K	3.7~4.8	4.1	4.1	3.9
Cl	101~109	104	104	102
Ca	8.7~10.1	8.9	8.8	9.4
無機リン	2.8~4.6	2.8	2.9	3.4
溶血 (H)	項目説明参照	7	5	6
黄疸 (I)	項目説明参照	1.5	1.8	2.5
乳び (T)	項目説明参照	1	1	1.8
血算				2
白血球数	3.4~9.2	5.9	4.8	6.1
赤血球数	4.00~5.66	4.70	4.64	4.86
ヘモグロビン	13.0~17.0	15.5	15.5	15.9
ヘマトクリット	38.2~50.8	45.4	45.3	47.1
MCV	85~102	96.6	97.6	96.9
MCH	28.4~34.6	33.0	33.4	32.7
MCHC	32.5~35.5	34.1	34.2	33.8
血小板数	141~327	178	155	162
比重	1.002~1.030	1.011	1.020	1.024
尿定性				1.013
pH (弱酸性)	5.5	5.5	5.5	5.5
糖	-	-	-	-
蛋白	-	-	-	+
潜血	-	-	-	H
ケトン体	-	-	-	-
ビリルビン	-	-	-	-
ウロビリノーゲン	1+	1+	1+	1+
亜硝酸塩	-	-	-	-
WBC定性	-	-	-	-
色調	淡黄色	淡黄色	黄色	淡黄色
濁度	-	-	-	-

# パーソナルゲノム

親から受け  
継いだゲノム

がんを発症  
する鍵遺伝  
子の変化

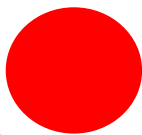
環境因子に  
よる変化を  
受けたゲノム

様々なリスクがわかる。治療方針が立つ。

これまでは「研究」のために行っていたシーケンスに対し、  
がんや患者さんの全DNA情報を  
シーケンスし、臨床的に翻訳・解  
釈することが現実になった。



# 臨床シーケンスの時代に入った

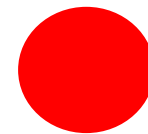


全ゲノム臨床シーケンス(患者さんに返す)研究がはじまった。

- 2009-2011年: [“One In A Billion”](http://www.pulitzer.org/archives/9180) (<http://www.pulitzer.org/archives/9180>): 米国ウイスコンシン医科大学が世界で初めての[全遺伝子解析に基づく治療](http://www.genome.gov/Multimedia/Slides/GenomicMedicineII/GMII_HJacob_ClinicalSequencingMCW.pdf)を行った。  
[http://www.genome.gov/Multimedia/Slides/GenomicMedicineII/GMII\\_HJacob\\_ClinicalSequencingMCW.pdf](http://www.genome.gov/Multimedia/Slides/GenomicMedicineII/GMII_HJacob_ClinicalSequencingMCW.pdf)  
そのレポート記事に2011年Pulitzer賞が与えられる。<http://www.pulitzer.org/archives/9180>
  - 米国NIH所長Francis Collinsが[2012年予算要求演説](http://www.nih.gov/about/director/budgetrequest/fy2012budgetrequest.pdf)の結論の中で引用。  
<http://www.nih.gov/about/director/budgetrequest/fy2012budgetrequest.pdf>
- 2011年12月6日: 米国NIH発表 “ゲノムシーケンスに基づく医療応用研究に\$4.16億ドル”(4年間で400億円)
- 2012年2月2日: カナダ [Ontario Institute for Cancer Research](http://oicr.on.ca/files/public/OICR_NR_February_02_2012.pdf) は、国際がんゲノムコンソーシアムの成果に基づき、がんの臨床シーケンスに基づく個別化医療の研究開始を発表 [http://oicr.on.ca/files/public/OICR\\_NR\\_February\\_02\\_2012.pdf](http://oicr.on.ca/files/public/OICR_NR_February_02_2012.pdf)
- 2012年5月3日~4日: 米国NIHが主催した[Genomic Medicine III](http://www.genome.gov/27548693)では検査の提供側、医療保険会社、クラウド会社が出席し、国を挙げての取り組み  
「米国では明日の問題！」 <http://www.genome.gov/27548693>
- 2012年6月20日: トロントの [Hospital for Sick Children](http://www.bio-itworld.com/news/06/20/12/Toronto-Hospital-Sick-Children-selects-Ion-Proton-whole-genome-sequencing.html) が半導体シーケンサー(Ion Proton)を導入し、将来的には、1万人/年の規模で全ゲノム臨床シーケンスをすることを発表(<http://www.bio-itworld.com/news/06/20/12/Toronto-Hospital-Sick-Children-selects-Ion-Proton-whole-genome-sequencing.html>)
- 米国臨床ゲノム会議: [The Clinical Genome Conference, San Francisco. June 11-13, 2012](http://www.genome.gov/27548693) 次回は2013年6月25-28(サンフランシスコ)
- 2012年7月8日: [50,000人、島全体シーケンス](http://sciencenordic.com/sequencing-genome-entire-population) FarGen Project  
<http://sciencenordic.com/sequencing-genome-entire-population>



# 臨床シーケンスの時代に入った



全ゲノム臨床シーケンス(患者さんに返す)研究がはじまった。

- 2011年12月6日: 米国NIH発表 “ゲノムシーケンスに基づく医療応用研究に\$4.16億ドル”(4年間で400億円)
- 2012年2月2日: カナダ [Ontario Institute for Cancer Research](#) は、国際がんゲノムコンソーシアムの成果に基づき、がんの臨床シーケンスに基づく個別化医療の研究開始を発表
- 2013年1月28-29日 ~ 2011年6月29日: 米国NIHが主催したGenomic Medicineの会議が4回クローズドで開かれ(80名ほど)、個別化ゲノム医療センターのほか、検査の提供側、医療保険会社、クラウド会社、医師のゲノム医療教育など、国を挙げて取り組んでおり、今年9月には “[Genomic Medicine is Global](#)” をテーマに第6回目が開催され、国際連携が議論される予定。

「米国ではまさに明日の問題！」



Dr. Lukas Wartman  
Washington Univ. St. Louis  
白血病の研究者が白血病にな  
り...

次世代シーケンサーとスパコ  
ンによるデータ解析で、腎癌の  
抗癌剤“スーテント”の有効性  
が示唆され、彼の白血病が寛  
解

"I was definitely scared. It was so unreal," said Dr. Wartman on first suspecting that he had leukemia, the very disease he had devoted his medical career to studying.

A drug that had been tested and approved only for advanced [kidney cancer](#). Dr. Wartman became the first person ever to take it for leukemia.

<http://www.nytimes.com/2012/07/08/health/in-gene-sequencing-treatment-for-leukemia-glimpses-of-the-future.html>

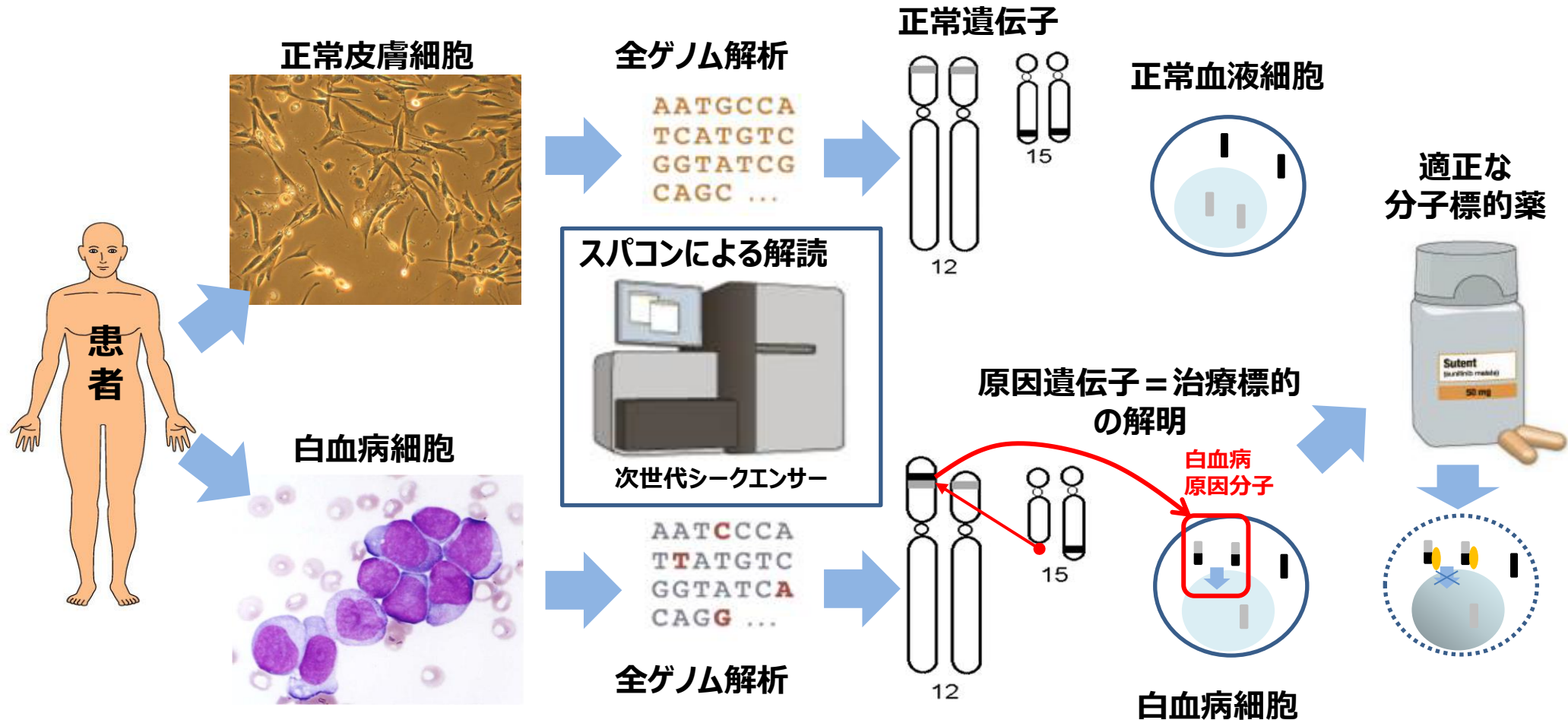
# メディカルインフォマティクスにもとづく白血病の個別化医療

The New York Times

2012年7月8日

Genetic Gamble

~New Approaches to Fighting Cancer



# 個別化医療のためのメディカルインフォマティクス(がんを例として)

全ゲノム解析をしていない従来の医療では

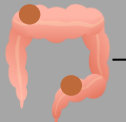
大腸がん患者さんでは、肺がんの原因遺伝子の変異は無視される



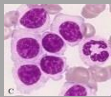
- Tamoxifen
- Anastrozole
- Doxorubicin
- Paclitaxel



- Cisplatin
- Etoposide
- Amrubicin
- Cyclophosphamide
- Doxorubicin



- 5FU
- UFT
- Oxaliplatin
- Irinotecan



- Idarubicin
- Daunorubicine
- Cytarabine

臓器別による抗癌剤の選択のみ



最先端シーケンサー



全ゲノム解析

AATGCCA  
TCATGTC  
GGTATCG  
CAGC ...

薬剤応答システム解析



パーソナルゲノム・トランスクリプトーム・エピゲノム・  
プロテオーム・メタボローム

個別化医療のための  
メディカルインフォマティクス

翻訳・  
解釈

バイオメディカル  
ビッグデータ解析

大規模のゲノム及び遺伝子発現  
データ、薬理・疾患情報のデータ  
解析による個別化医療の基礎データ・知識ベース

全ゲノム・全トランスクリプトーム解析をすると

私のがんにぴったりあった副作用の無い抗癌剤と治療法がありました!



ゲノム変異別分類

抗がん剤・治療法

HER2 amplification

- Trastuzumab
- Lapatinib

EGFR mutation

- Gefitinib
- Erlotinib

ALK fusion

- Crizotinib

RET fusion

- Vandetanib

BRAF mutation

- Vemurafenib

PIK3CA mutation

- Temozolomide

BCR-ABL fusion  
cKIT mutation

- Imatinib

*Open*[Geisinger Health System](#)

## Implementing genomic medicine in the clinic: the future is here

Teri A. Manolio, MD, PhD<sup>1</sup>, Rex L. Chisholm, PhD<sup>2</sup>, Brad Ozenberger, PhD<sup>1</sup>, Dan M. Roden, MD<sup>3</sup>,  
Marc S. Williams, MD<sup>4,5</sup>, Richard Wilson, PhD<sup>6</sup>, David Bick, MD<sup>7</sup>, Erwin P. Bottinger, MD<sup>8</sup>,

Although the potential for genomics to contribute to clinical care has long been anticipated, the pace of defining the risks and benefits of incorporating genomic findings into medical practice has been relatively slow. Several institutions have recently begun genomic medicine programs, encountering many of the same obstacles and developing the same solutions, often independently. Recognizing that successful early experiences can inform subsequent efforts, the National Human Genome Research Institute brought together a number of these groups to describe their ongoing projects and challenges, identify common infrastructure and research needs, and outline an implementation framework for investigating and introducing similar programs elsewhere. Chief among the challenges were limited evidence and consensus on which genomic variants were medically relevant; lack of reimbursement for genomically driven interventions; and burden to patients and clinicians of assaying, reporting, intervening, and following up genomic findings. Key infrastructure needs included an openly accessible knowledge base capturing sequence variants and their phenotypic associations and a framework for defining and cataloging clinically actionable variants. Multiple institutions are actively engaged in using genomic information in clinical care. Much of this work is being done in isolation and would benefit from more structured collaboration and sharing of best practices. *Genet Med* advance online publication 10 January 2013 *Genetics in Medicine* (2012); doi:10.1038/gim.2012.157.



# Mayo Clinic



Richard Weinshilboum教授(薬理ゲノム)

- メイヨーは5年以内に10万人の全ゲノム情報のデータを取り、電子カルテとリンクさせる
- 専用のコンピュータ棟も建築中

(2013年1月の情報)

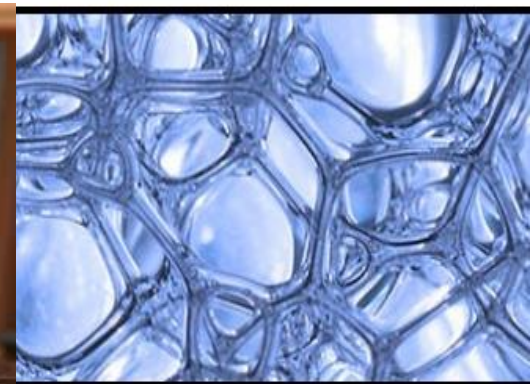
メイヨークリニックのあるミネソタ州ロチェスターは人口10万人の町で、メイヨークリニック関係の雇用が3万5千人、医師がレジデントとを含めると5000人。入院ベッドは2000床で、年間延べ患者数は100万人、この内隣接する地域が50%、その他の米国からが45%、海外から5%(5万人)。ロチェスター空港はボーディングブリッジは2つしかありませんが、著名人がメイヨーを受診するため、プライベートジェットで乗り付けるそうです。完全な医療都市です。アリゾナとフロリダにも病院を持っています。

# シカゴ大学 [Grossman](#) 教授のブログ

## 100万人ゲノムが医療を変える

- **Managing and Analyzing 1,000,000 Genomes:** September 18, 2012 by Robert Grossman
- Recently, we have been thinking about what you might call the [Million Genome Challenge](#). Over the next several years, the National Cancer Institute, and perhaps other organizations, will sequence a million genomes and use this data to increase our understanding of biological pathways and of genomic variation across individuals. [With this knowledge, we can begin to stratify diseases, leading to precision diagnosis and precision treatment that is personalized for individual patients.](#)
- [The numbers associated with a million cancer genomes are worth thinking about.](#) The whole genome data for a tumor and a matching normal tissue sample require about 1 TB. Thus, one million genomes require about 1,000,000 TB. This is 1,000 PB or 1 EB. Compressing the data might reduce the data by about a factor of 10. Throwing away the alignment data and retaining only the variation data would reduce the data by about a factor of about 100. Assuming it costs about \$1,000 to sequence each whole genome, the project as a whole requires about \$1B for the sequencing. It might require another \$1B for the infrastructure and analysis. Although obviously a large project, [a project like this is likely to fundamentally alter the way we understand and treat diseases.](#)

### Robert Grossman



[Group](#) [Software](#) [Blog](#) [Contacting](#)

← [Why is it still so hard to analyze remote and distributed data?](#)

[Datascope for the Long Tail of Science](#) →

### Managing and Analyzing 1,000,000 Genomes

Posted on [September 18, 2012](#) by [Robert Grossman](#)

I gave a talk last week at [XLDB 2012](#) about [Bionimbus](#), which is cloud based system for managing, analyzing, transporting, and sharing large genomics datasets in a secure and compliant fashion. Bionimbus was developed at the Institute for Genomics and Systems Biology ([IGSB](#)) and is used by IGSB and some of their collaborators to manage and analyze their next gen sequencing data.



## Knocking on the clinic door

*Nature Biotechnology* **30**, 1009 (2012) | doi:10.1038/nbt.2428

Published online 08 November 2012

[Download PDF](#)[Citation](#)[Reprints](#)[Rights & permissions](#)

**High-throughput sequencing for clinical purposes faces technical and quality challenges, but it's worth it.**

The \$3-billion price tag for the first human genome would now buy not one but a million human genome sequences, each completed in just a few weeks. Personal genome sequencing is becoming a reality, and targeted or whole exome sequencing is being explored to facilitate diagnosis and guide treatment, in some conditions and for some patients. The problem is that extracting clinically actionable information from genome data is currently hit or miss, time intensive and dependent on access to knowledgeable specialists. What's more, much of the IT infrastructure and decision support systems necessary to deliver genome information to physicians has yet to be put in place.

# 我が国の風景

# パーソナルゲノムに基づく個別化医療の推進(これからの医科研5年間)

がん・感染症  
患者



2001年より遺伝子カウンセリングと遺  
伝子検査体制が整備されている。

医科研附属病院

お金はないですが、  
やります！



遺伝カウンセリング

検体

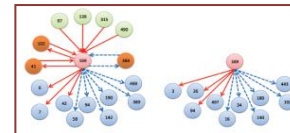
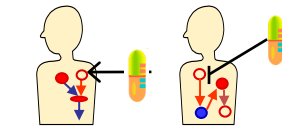
超並列シーケンサー  
2013年ヒトゲノムは10万円以下で解析

スパコンによるデータ解析  
225 TFLOPS計算ノード

全DNA解析  
トランスクリプトーム解析  
エピゲノム解析

パーソナルオミクス情報  
1人当たり0.5TBを超えるデータ

医療介入の予測



- 最適な治療法・  
治療薬量の決定
- 副作用の回避
- 個別化検診・  
サーベイランス  
の実施
- 体質に応じた予  
防法

Human Genome Center  
Institute of Medical Science, University of Tokyo



4.4PBストレージ  
3PB Lustre File System



ネットワーク分離  
スパコン  
高セキュリティ  
データ解析室



先端医療研究センター・医科研附属病院

ヒトゲノム解析センター

スパコン: 国際がんゲノムコンソーシアムを初めとするゲノムシーケンスデータ解析の体制が整っている。

血液がん: 年間400件以上の多様な遺  
伝子検査と診療の実績  
大腸がん: 100症例以上の遺伝性大腸  
がんの遺伝子診断と診療

Personalized  
Medicine  
Team

# 黒船来航 1853年

「DNA二重らせんの発見1953年のちょうど100年前、  
黒船が日本を覚醒させましたが・・・」  
2003年のヒトゲノム解読から10年後の今・・・



THE ARRIVAL OF THE AMERICAN SQUADRON IN EDO BAY, 1853. BY G. H. PHILLIPS.

TO VIEW THE AMERICAN COMMISSIONERS IN THE SQUADRON, JOHN WALKER & CO.

10, NASSAU ST. N.Y. (SEE ADVERTISING COLUMN FOR PARTICULARS.)

# “苦悶”

## ヒトゲノム解析センター

- 計算能力: 225 TFLOPS (ピーク性能)
- ストレージ: Lustre 高速ストレージ(3PB) + ニアライン(2PB)
- リース: 6年リースで60億円弱 (2009.1~2014.12)
- 消費電力: 実測値550KW~800KW (ピーク値1200KW)
- 多様なライフサイエンスアプリに対応: **22,704コア**

2012年11月18日  
世界300位

2009年6月26日  
世界69位  
ライフサイエンス特化:  
世界2位

	ノード数	CPU	メモリ/ノード	総メモリ	総コア数
Shirokane 2(分散メモリサーバ)	504	CPU:AMD Opteron 6276 2.3GHz (16cores) × 2	32GB	16.128TB	16,128
Shirokane2(分散メモリサーバ)	12	CPU:Intel Xeon X5675 3.06GHz (6cores) × 2	144GB	1.728TB	144
Shirokane2(共有メモリサーバ)	1	CPU:Intel Xeon E7 8837 2.66GHz (8cores) × 16	2,000GB	2TB	128
Shirokane1(分散メモリサーバ)	764	CPU:Intel Quad Core Xeon E5450 3.0GHz (4cores) × 2	32GB	24.448TB	6,112
Shirokane1(分散メモリサーバ)	12	CPU:AMD Quad Core Opteron 8356 2.3GHz (4cores) × 4	128GB	1.536TB	192

# SHIROKANE システム全体構成図

3階

共有メモリ型サーバ

分散メモリ型サーバタイプA

分散メモリ型サーバタイプB

分散メモリ型サーバタイプB

分散メモリ型サーバタイプA

データバックアップサーバ

高速ディスクアレイ

CPU: Intel Xeon E7 8837 2.66GHz (8cores) × 16  
メモリ: 2TB

12ノード  
CPU: AMD Quad Core Opteron 8356 2.3GHz (4cores) × 4  
メモリ: 128GB

764ノード  
CPU: Intel Quad Core Xeon E5450 3.0GHz (4cores) × 2  
メモリ: 32GB

504ノード:  
CPU: AMD Opteron 6276 2.3GHz (16cores) × 2  
メモリ: 32GB

12ノード  
CPU: Intel Xeon X5675 3.06GHz (6cores) × 2  
メモリ: 144GB

1階

容量 500TB

データバックアップサーバ

高速データ通信ネットワーク

高速データ通信ネットワーク

容量 500TB

容量 1PB

容量 500TB

容量 2.3PB

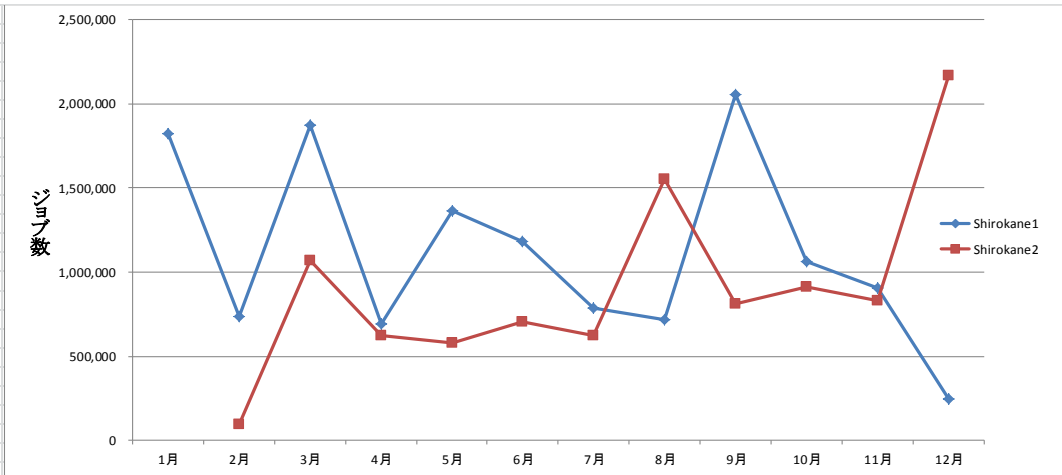
DBサーバ

# パーソナルゲノム時代の 基盤ファイルシステム

ヒトゲノム解析センターの場合  
ライフサイエンス系利用に特化したコンピュータ

# メタデータアクセス

- ジョブの数、ジョブの入出力ファイルの数が多い
- 1ジョブセットで数万ファイルが作成・参照されることも多い



ヒトゲノム解析センター	2012年度実績
Shirokane1	13,446,923 ジョブ
Shirokane2	9,968,079 ジョブ
合計	23,415,102 ジョブ

比較のためにつけております

東大情報基盤センター	2012年年度実績
FX10	322,555 ジョブ (4月-12月)
HA8000	91,979 ジョブ
SRI16000	58,041 ジョブ



# 少し昔を懐古してみると 国際ヒトゲノム計画

- 1953年：ワトソンとクリックによるDNAの二重らせん構造の発見



- 2003年にこの30億文字からなるゲノム情報の電子化を完了しました。
- 「ヒトゲノムが解読された」と世界中で報道されましたが、ヒトゲノムという暗号が解読されたわけではありません。
- 13年の歳月と何もかもあわせて1000億円かかりました。日本の京の開発より少額。

2003年～2013年  
今日までの10年

失われた10年



# 米国NIH ヒトゲノム解読後のロードマップ 2003

***Biology*** is changing fast into a  
***Science of Information Management.***”

ICTによる医療・ヘルスケア開発のパラダイムシフト  
が起こることが示唆されていました。

現在、このパラダイムシフトは急速に進行しています。

# 2003～2007年 国際ハブマップ計画

## IN SCIENCE Breakthrough of the Year 2007

### Editorial: Breakthrough of the Year >

Science Editor-in-Chief Donald Kennedy overviews the big stories from 2007 covered in this year's Breakthrough issue.

### Breakthrough of the Year: Human Genetic Variation >

Equipped with faster, cheaper technologies for sequencing DNA and assessing variation in genomes on scales ranging from one to millions of bases, researchers are finding out how truly different we are from one another.

### It's All About Me >

Along with the flood of discoveries in human genetics, 2007 saw the birth of a new industry: personal genomics. But researchers worry that these services open up a Pandora's box of ethical issues.

### The Runners-Up >

The runners-up for 2007's Breakthrough of the Year include advances in cellular and structural biology, astrophysics, physics, immunology, synthetic chemistry, neuroscience, and computer science.

### Scorecard: How'd We Do? >

Some of last year's predictions panned out this year, especially the work that led to the Breakthrough of the Year, but other areas are progressing more slowly.

Get More from  
Science with a Free  
Access Account



Watch a [video presentation](#) on this year's discoveries in human genetic variation, featuring Francis Collins, David Altshuler, and *Science* news writer Liz Pennisi.

> [Higher-bandwidth version of video](#)

> [Lower-bandwidth version of video](#)

- 個人個人の違いをDNAのレベルで解明する国際ハブマップ計画が完了し、ヒトの病気や薬剤応答に関わる遺伝子を効率よく見つけるための基盤ができました。
- 3つの人種の全ゲノムにわたるDNAの違いが明らかにされました。

# 米国 2008年～2010年

- 2008年、「ゲノムと個別化医療法案」を提出したオバマ上院議員が米国大統領に就任しました。
- 2009年、「**遺伝子差別禁止法**」が成立しました。
- 2010年、

## The Genomics and Personalized Medicine Act Returns to Congress

Posted by [Guest Contributor](#) and [Dan Vorhaus](#) on August 24, 2010



Meggan Bushee is a student at the [Wake Forest University School of Law](#).

This past May, Congressman Patrick Kennedy (D-RI) and Congresswoman Anna Eshoo (D-CA) re-introduced a personalized medicine bill to the U.S. House of Representatives. The bill was originally introduced in 2006 by then-Senator from Illinois Barack Obama. While HR 5440, also known as the [Genomics and Personalized Medicine Act of 2010](#) (GPMA 2010), has retained the name of the bill originally introduced by Senator Obama, its approach to the regulation of personalized medicine has taken a new direction.

GPMA 2010 is the fourth version of the GPMA since the original bill of 2006, and includes the most ambitious initiatives of all of its predecessors. Why has the GPMA re-surfaced after three prior versions failed to make it out of committee? According to Representative Kennedy, the bill has been re-introduced in response to increased public awareness and use of genomic tests. At present, GPMA 2010 is before the House Committee on Energy and Commerce. This is the same committee that [recently conducted high-profile hearings to review the current state of the direct-to-consumer \(DTC\) genetic testing registry](#).



# 2008年～国際がんゲノム計画

- 主要ながんのゲノム異常カタログを作成する。
- 25,000のがんサンプルと正常細胞、あわせて**5万人**分の全ゲノム情報が解析されています。

Bladder Cancer United States	Blood Cancer United States	Bone Cancer United Kingdom
Brain Cancer United States	Breast Cancer European Union / United Kingdom	Breast Cancer France
Breast Cancer Mexico	Breast Cancer United Kingdom	Breast Cancer United States
Cervical Cancer United States	Chronic Lymphocytic Leukemia Spain	Chronic Myeloid Disorders United Kingdom
Colon Cancer United States	Endocrine Tissues Cancer No jurisdiction(s) committed	Endometrial Cancer United States
Esophageal Cancer United Kingdom	Gall Bladder & Biliary System Cancer No jurisdiction(s) committed	Gastric Cancer China
Gastric Cancer United States	Head and Neck Cancer United States	Liver Cancer France
Liver Cancer Japan	Liver Cancer United States	Lung Cancer United States
Malignant Lymphoma Germany	Oral Cancer India	Ovarian Cancer Australia
Ovarian Cancer United States	Pancreatic Cancer Australia	Pancreatic Cancer Canada
Pediatric Brain Tumors Germany	Prostate Cancer Canada	Prostate Cancer Germany
Prostate Cancer United Kingdom	Prostate Cancer United States	Rare Pancreatic Tumors Italy
Rectal Cancer United States	Renal Cancer European Union / France	Renal Cancer United States
Skin Cancer United States	Soft Tissue Cancer No jurisdiction(s) committed	

[Launch Data Portal »](#)

[Apply for Access to Controlled Data »](#)

## Announcements

5/June/2011 - The ICGC Data Coordination Center is pleased to announce the release of version 5 of the [ICGC data portal](#). This release includes the first data submission from the Chronic Lymphocytic Leukaemia (CLL) project, whose work was published in Nature on June 5th ([Read the article](#)). The addition of this new project brings the total number of cancer projects in the release to 25. With this release, the CLL project also becomes the first to host its own ICGC data portal server. Users visiting the ICGC data portal now have the option of choosing to access all ICGC datasets through the portal hosted in Canada at OICR, or in Spain at the Barcelona Supercomputing Center (BSC); this feature is available through a dropdown menu at the top of the portal page.

## Updates

Currently, the ICGC has received commitments from funding organizations in Asia, Australia, Europe and North America for 39 project teams in 13 jurisdictions to study over 18,000 tumor genomes. Projects that are currently funded are examining tumors affecting the bladder, blood, bone, brain, breast, cervix, colon, head and neck, kidney, liver, lung, oral cavity, ovary, pancreas, prostate, rectum, skin, soft tissues, stomach and uterus. Over time, additional nations and organizations are anticipated to join the ICGC. The genomic analyses of tumors conducted by ICGC members in Australia and Canada (pancreatic cancer), Japan (liver cancer), Spain (blood cancer), the UK (breast, lung and skin cancer) and the USA (blood, brain, breast, colon, kidney, lung, ovarian, rectal, stomach and uterine cancer) are now available through the Data Coordination Center housed on the ICGC website at [www.icgc.org](http://www.icgc.org).

International network of cancer genome projects. Nature 464, 993-998 (15 April 2010)

[Read the article »](#)



## About ICGC

The International Cancer Genome Consortium (ICGC) has been organized to launch and coordinate a large number

# 2009年 米国NIH

- 2009年、米国ヒトゲノム計画を完遂したコリンズ博士が米国National Institute of Health所長に就任。ゲノム情報を基盤とした米国の医療・ヘルスケア戦略が、だれの目にも明らかになりました。



# 2010年 中・英・米

- 中・英・米は、世界各地の1000人以上の全ゲノムを解析して、医学的応用価値のある人類ゲノム地図を作成する「1000人ゲノム」のゲノムデータができました。
- 人類のゲノム地図が完成する。
- データは次のサイトからダウンロードできます。

<http://www.1000genomes.org/>



困難はいろいろありますが

治療の前に

あなたのがんのゲノムを調べますか？

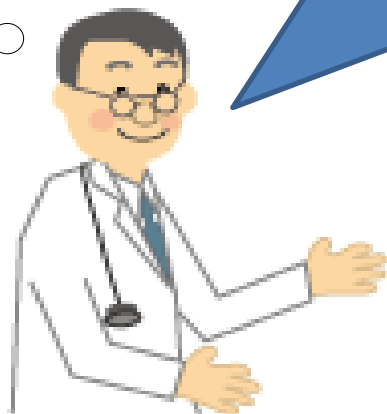
調べないで治療しますか？



未来はとっくにはじまっている。

- あなたのがんには、遺伝子AとBの異常がありました。
- あなたには抗癌剤XとYが有効です。
- 抗癌剤Y使うと副作用があります。
- あなたは遺伝性のがんではありません。

あなたのがんにぴったり合った薬があって良かったですね。



だれもが保険証を持っているように自分のDNA情報をもって医療を受ける時代がやがて来る。

## DNA情報に基づいた個別化医療

### 遺伝子 医療革命

ゲノム科学がわたしたちを変える  
The Language of Life

DNA and the Revolution in Personalized Medicine

フランシス・S・コリンズ Francis S. Collins 矢野真千子訳

史上最強のアイスホッケー選手とだれもが認めるウェイン・グレツキーは、彼の最初にして最高のコーチである父親ウォルターから、「バックが行きそうなところに行け」というシンプルな教えを受けていたという。私たちはみな、人生というアイスホッケー試合のスケーターだ。チームのメンバーと協力しながら、転ばないように注意しながら、そしてできることなら数回ゴールを決めながら、リンクの上を巧みに動き回ろうとしている。そのためには、バックを追いかけるだけではだめだ。バックが行きそうなところに行っておかなければならない。あなたのDNAの二重らせん、あなたの生命の言語は、あなただけの医学書だ。その読み方を学び、人生に活かすことができるのは、あなたしかいない。

Francis S. Collins著 (米国National Institute of Health所長)  
(邦訳 NHK出版)

