

Tackling Biomedical Big Data

東京大学医科学研究所ヒトゲノム解析センター

宮野 悟

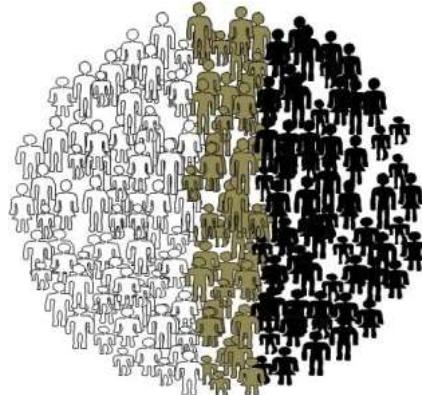
将来HPCIあり方調査研究「アプリ分野」第6回全体ミーティング

2013年5月31日, TKP東京駅ビジネスセンター1号館 3階 ホール3A

「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

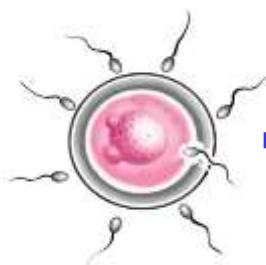


「がんは日本の国民病」日本人の半分が罹り、3分の1が亡くなっている



「私のDNA」と「私」

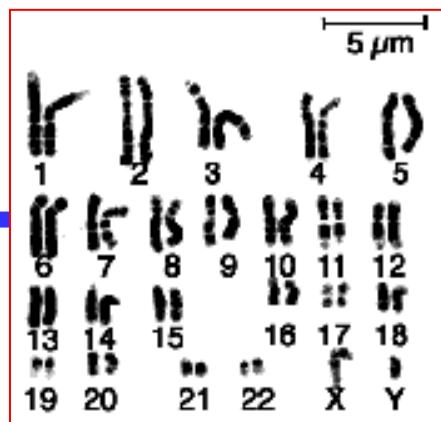
母のDNAと卵細胞



1954年

父のDNA

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
病気
ボケ、がん



死

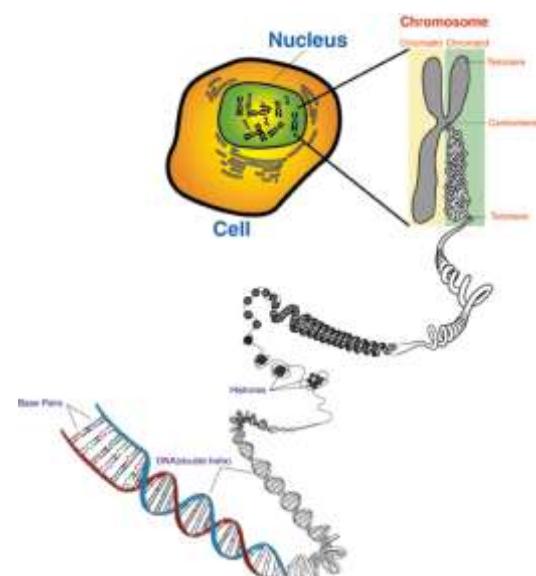


2013年

皆さん的人生(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

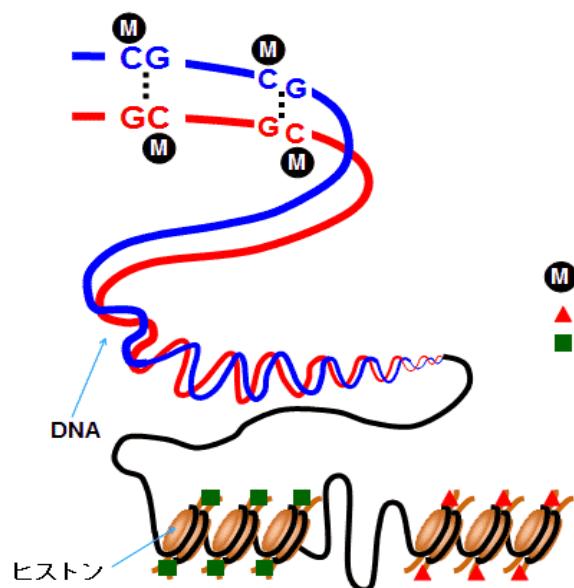
DNA

ゲノム



DNAやヒストンの修
飾状態

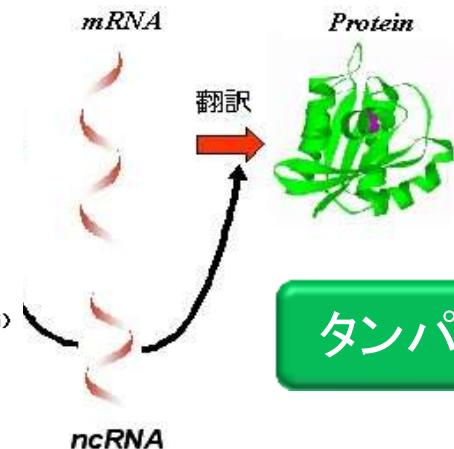
エピゲノム



遺伝子発現データ

mRNAやnoncoding
RNA

トランскriプトーム



タンパク質

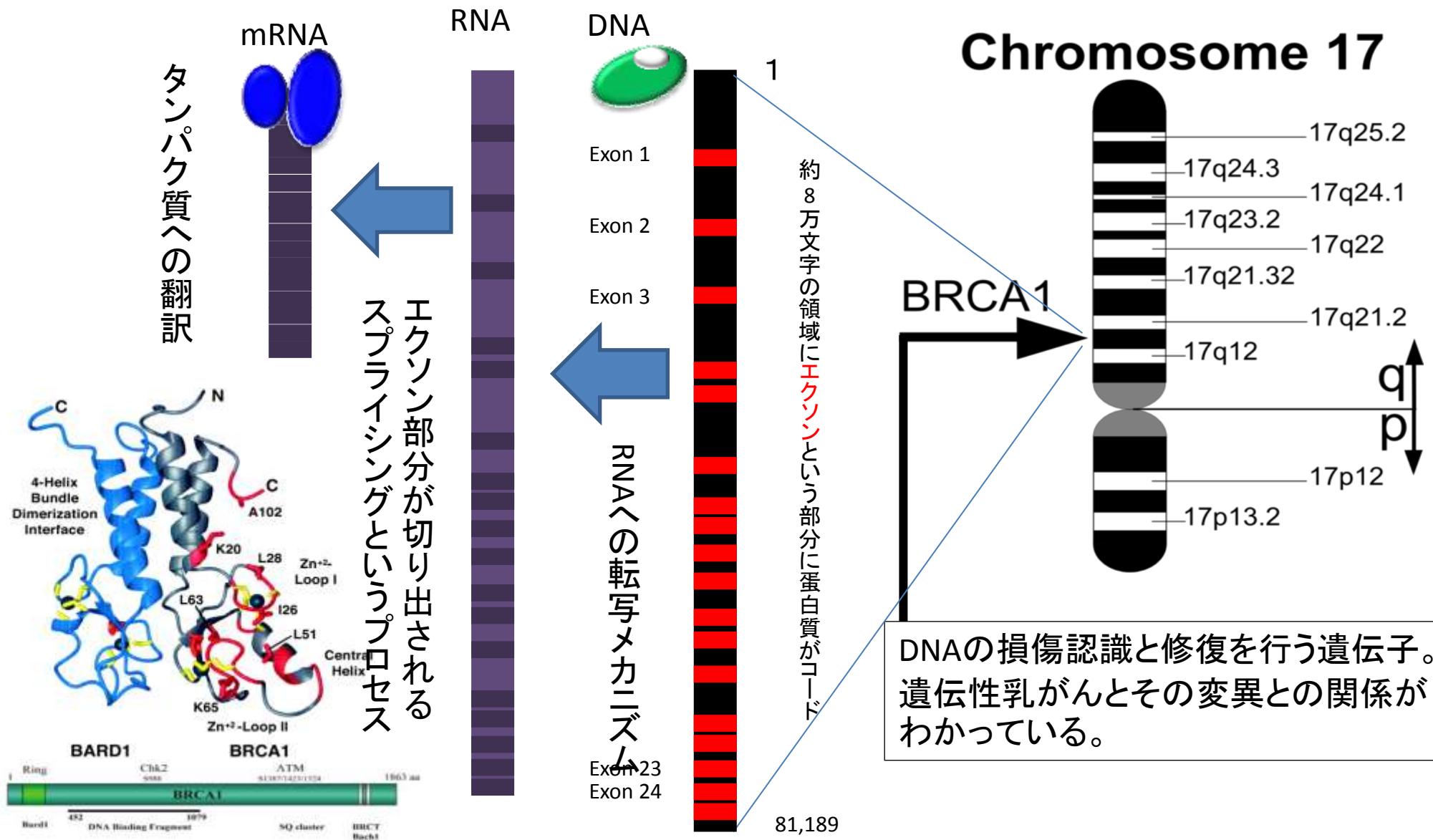
タンパク質に翻訳されない
DNAからの転写産物

Image source:
www.biotec.or.th/Genome/whatGenome.html

分子生物学のセントラルドグマ

ヒト17番染色体

約8100万文字のDNA情報



THE WALL STREET JOURNAL.

日本 ワールド・ストリート・ジャーナル 2013年 5月 14日 (火)

有料会員登録受付中
1ヶ月購読で3ヶ月読み放題

有料会員登録

ホーム マーケット ビジネス 経済 テクノロジー 國際 国内 オピニオン ライフ バロンズ

TOP STORIES IN U.S.

1 of 12



Massive Online Courses Woo Professors



New Tack in Preventing Hospital Infections



How Three Small Firms Face Health Law



3 of 12

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

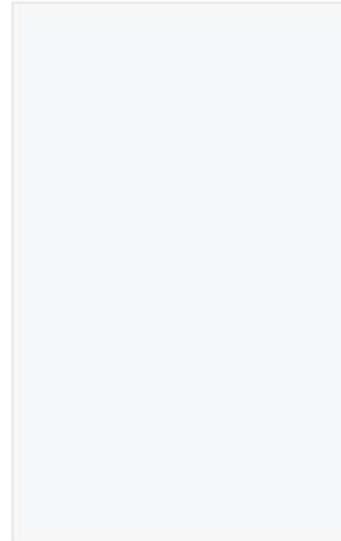
Angelina Jolie Says She Had Double Mastectomy

Article

Video

Comments (25)

Associated Press

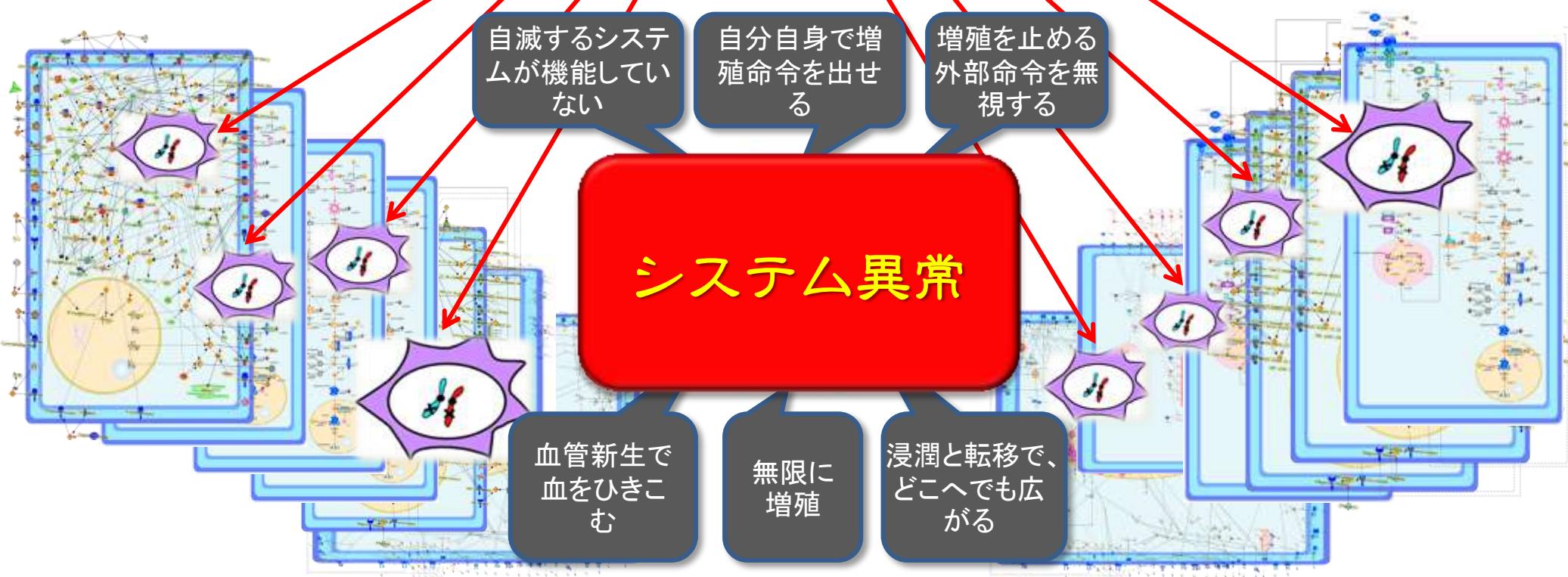
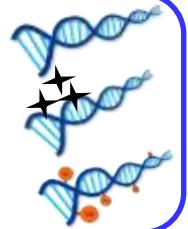


<http://www.youtube.com/watch?v=1qW3zwqp-kY>

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「シークエンス」

生物のDNA情報を読み取る装置は一般にシークエンサーとよばれ、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「シークエンス」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

★ シリコンシークエンサーの登場
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



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

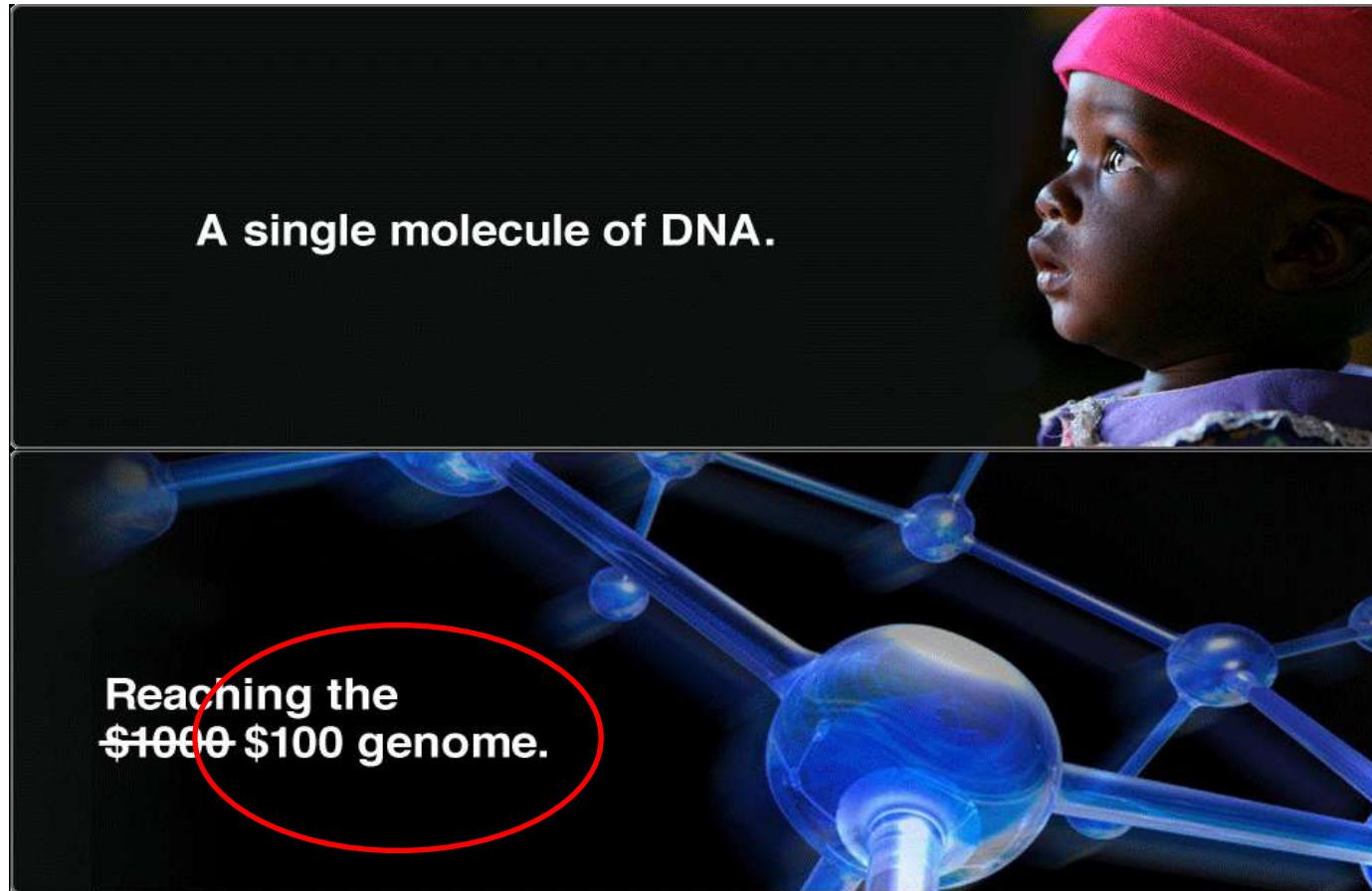
しかし、DNAビッグデータを解読して
活用できる人材・計算資源が不足



Genia Technologies, Inc.

2013年～ヒトゲノムは100ドル以下、1時間以内

“Last-Generation Sequencer”



「高速に大量のシークエ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

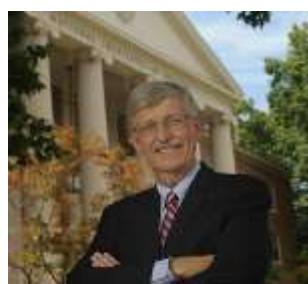
米国National Institute of Health所長

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-

11

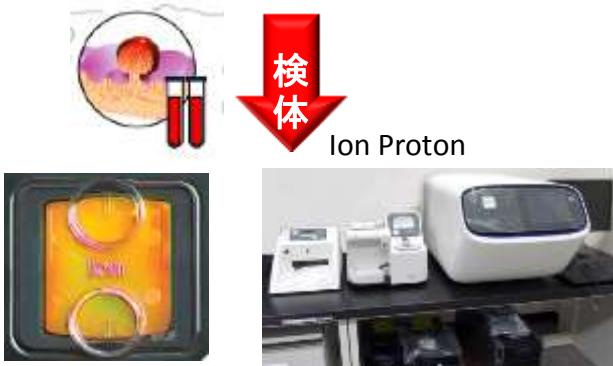


米国NIH所長
F. Collins

スーパーコンピュータで シークエンスのアセンブリとデータ解析



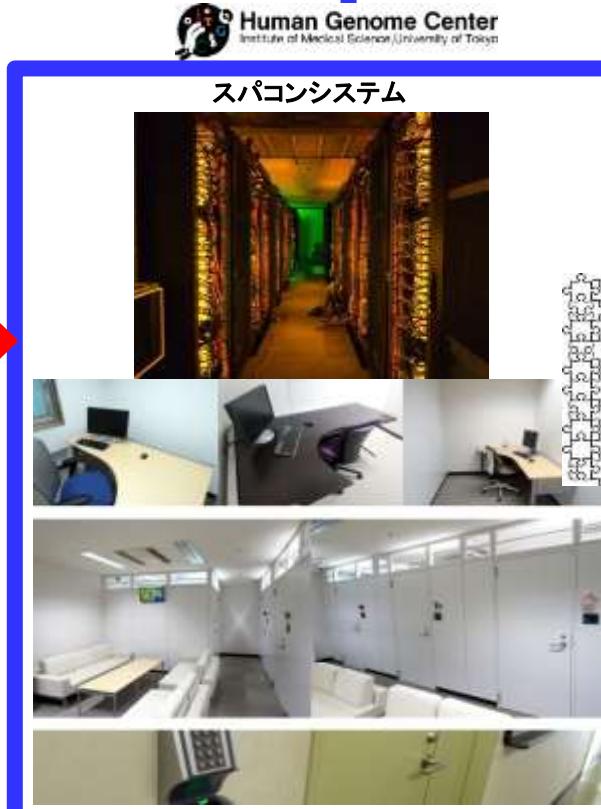
メディカルインフォマ
ティクスの達人たち



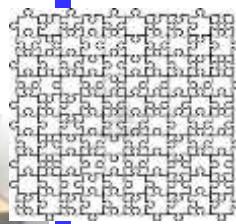
半導体シークエンサー



正常組織 900億文字
がん組織1200億文字
のデータ



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
による変化を
受けたゲノ
ム



*** 検査結果 時系列画面プリント ***

患者番号: [REDACTED] 氏名: [REDACTED]

プリント時間: 2012.4.9

13:23:35

P. I

採取日

2011.10.19

採取時間

00:00

地蛋白

6.9~8.4

TG^{アシニ}

3.9~5.2

CK

62~287

AST(GOT)

13~33

ALT(GPT)

8~42

LD

119~229

y GT

9~109

尿素窒素

8~21

クレチニ

0.6~1.1

eGFR

71.2

尿酸

2.6~7.0

空腹時血糖

70~110(空腹時)

食事日時

食事内容

採血日時

HbA1c(NGSP)

4.6~6.2

HbA1c(JDS)

4.3~5.8

中性脂肪

30~150

総コレステロール

122~240

HDLコレステロール

35~70

LDLコレステロール(直接法)

140未満

Na

139~146

K

3.7~4.8

Cl

101~109

Ca

8.7~10.1

無機リン

2.8~4.6

溶血(H)

2.8

項目説明参照

7

黄疸(I)

1.5

乳び(T)

1

血算

白血球数

3.4~9.2

赤血球数

4.00~5.66

ヘモグロビン

13.0~17.0

ヘマトクリット

38.2~50.8

MCV

85~102

MCH

28.4~34.6

MCHC

32.5~35.5

血小板数

141~327

比重

1.002~1.030

尿定性

pH

5.5

糖

-

蛋白

-

糖

-

タンパク

-

ヒメビン

-

リノバクタ

-

藍細胞菌

-

WBC定性

-

色調

-

濃度

-

淡黄色

淡黄色

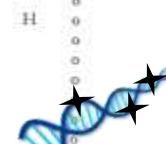
黄色

淡黄色

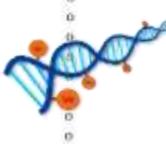
黄色

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

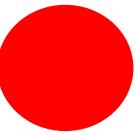
臨床シークエンスの時代に入った

全ゲノム臨床シークエンス(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を行った。
http://www.genome.gov/Multimedia/Slides/GenomicMedicine/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 “ゲノムシークエ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、がんの臨床シークエ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
- 2012年5月3日～4日: 米国NIHが主催した[Genomic Medicine III](#)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](#) が半導体シークエンサー(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クエ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](#) 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、島全体シークエンス](#) FarGen Project
<http://sciencenordic.com/sequencing-genome-entire-population>

臨床シークエンスの時代に入った



全ゲノム臨床シークエンス(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 “ゲノムシークエ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、がんの臨床シークエ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！」

2012年7月8日

Genetic Gamble

～New Approaches to Fighting Cancer



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>

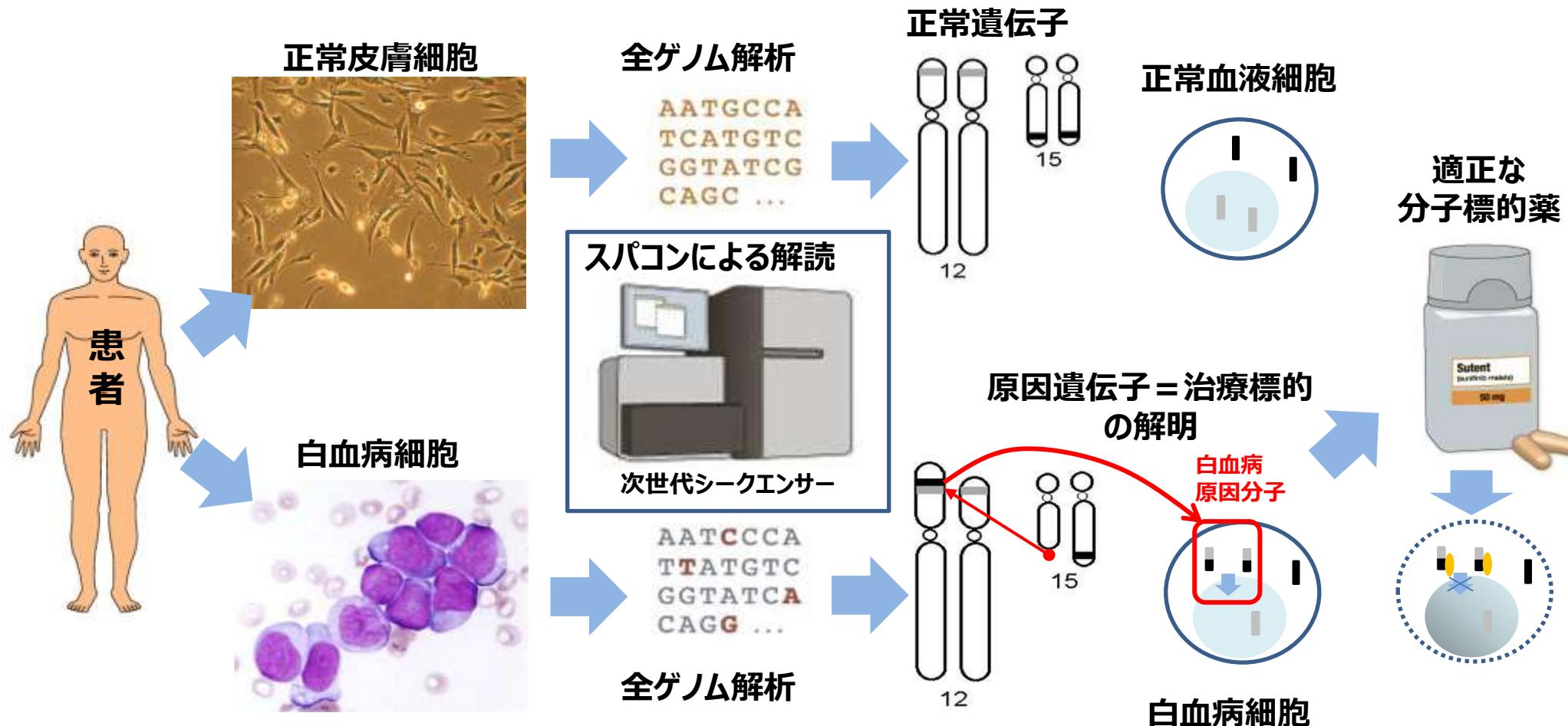
メディカルインフォマティクスにもとづく白血病の個別化医療

The New York Times

2012年7月8日

Genetic Gamble

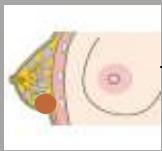
~New Approaches to Fighting Cancer



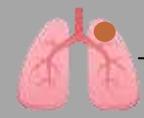
個別化医療のためのメディカルインフォマティクス(がんを例として)

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



全ゲノム・全トランск립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	抗癌剤・治療法
HER2 amplification	• Trastuzumab • Lapatinib
EGFR mutation	• Gefitinib • Erlotinib
ALK fusion	• Crizotinib
RET fusion	• Vandetanib
BRAF mutation	• Vemurafenib
PIK3CA mutation	• Temsirolimus
BCR-ABL fusion cKIT mutation	• Imatinib

Open

[Geisinger Health System](#)

Implementing genomic medicine in the clinic: the future is here

Teri A. Manolio, MD, PhD¹, Rex L. Chisholm, PhD², Brad Ozenberger, PhD¹, Dan M. Roden, MD³,
Marc S. Williams, MD^{4,5}, Richard Wilson, PhD⁶, David Bick, MD⁷, Erwin P. Bottinger, MD⁸,

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

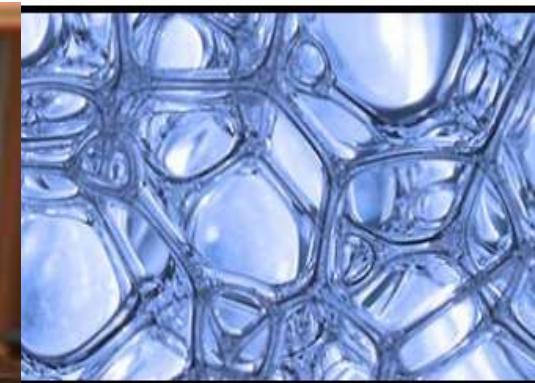
- **Managing and Analyzing 1,000,000 Genomes**

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?

Datascopes 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

超並列シークエンサー
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скriプトーム解析
エピゲノム解析

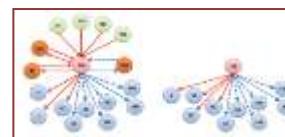
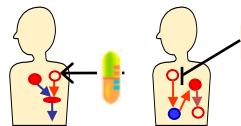


4.4PBストレージ
3PB Lustre File System

Human Genome Center
Institute of Medical Science, University of Tokyo



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

先端医療研究センター・医科研附属病院

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

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始めとするゲ
ノムシークエンスデータ解析の体制が整っている。

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 BLACK SHIPS AT EDO.
TO MEET THE AMERICAN COMMERCIAL SQUADRON, JAPAN, 1853.

“苦悶”

ヒトゲノム解析センター

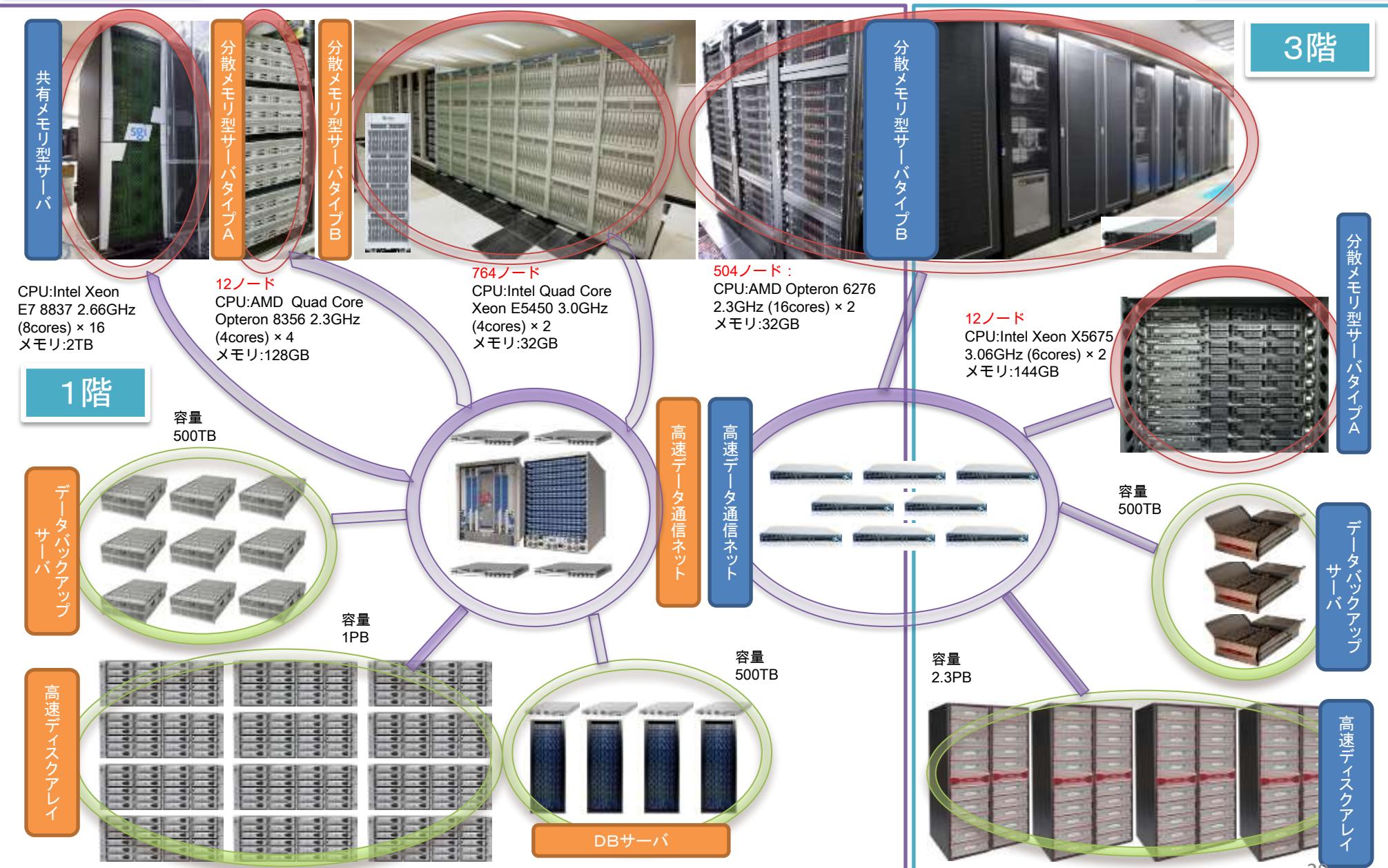
- 計算能力: 225 TFLOPS(ピーク性能)
- ストレージ: Lustre 高速ストレージ(3PB) + ニアライン(2PB)
- リース: 6年リースで60億円弱(2009.1~2014.12)
- 消費電力: 実測値550KW~800KW(ピーク値1200KW)
- 多様なライフサイエンスアプリに対応: [22,704コア](#)

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

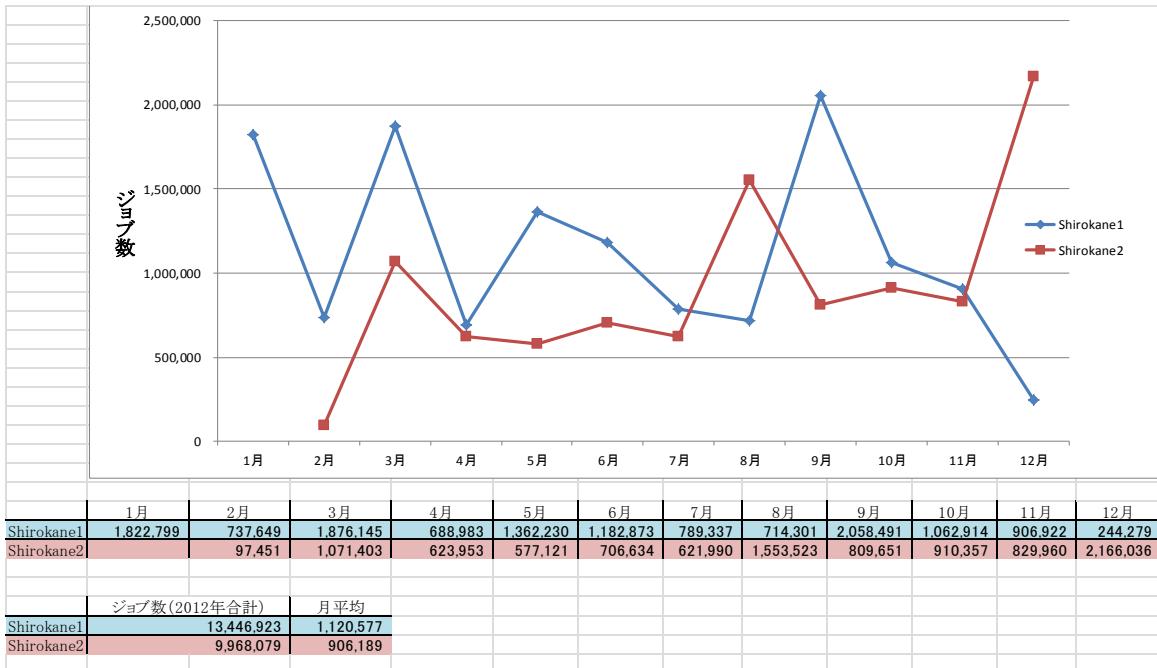


パーソナルゲノム時代の 基盤ファイルシステム

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



50年

- 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 国際ハプマップ計画

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.



- 個人個人の違いをDNAのレベルで解明する国際ハプマップ計画が完了し、ヒトの病気や薬剤応答に関す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 Buhshee 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 industry.



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。

[Launch Data Portal](#)

[Apply for Access to Curated Data](#)

[Announcements](#)

5 June 2011 - The ICGC Data Coordination Center is pleased to announce the release of version 5 of the [ICGC portal](#). This release includes the first data submission from the Chronic Lymphocytic Leukemia (CLL) project, whose work was published in *Nature* on June 16 ([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 CIHR, 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 29 project teams in 13 jurisdictions to study over 10,000 tumor patients. Projects that are currently funded are examining tumors affecting the bladder, blood, bone, brain, breast, 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 (breast 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](#).

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/>

スパコンで暴き出す がんのシステム異常の本態

スパコンをつかったがん研究のご紹介



新学術領域研究 システムがん (複合領域: 4201)

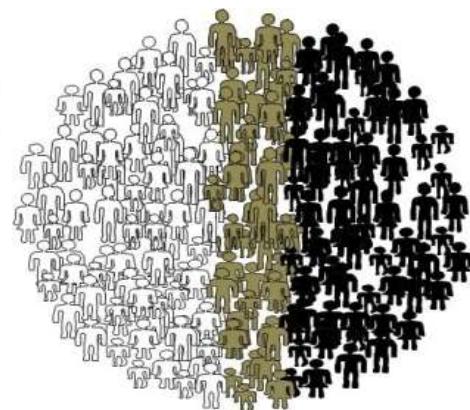
システム的統合理解に基づくがんの先端的診断、治療、予防法の開発

領域代表 宮野 悟 PhD 東大医科学研究所ヒトゲノム解析センター
A01班長 稲澤譲治 MD PhD 東京医科歯科大難治疾患研究所
A02班長 高橋 隆 MD PhD 名古屋大学医学研究科

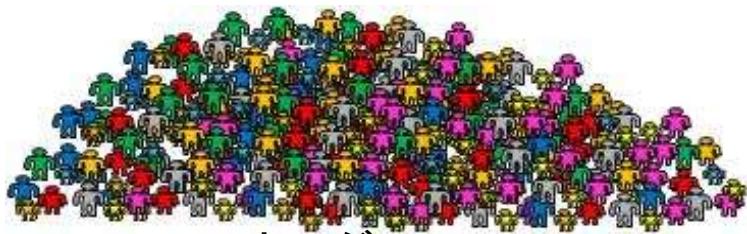
「世界に一つだけの花」
私たちは一人一人異なる
ゲノムを持っている



「がんは日本の国民病」日本
人の半分が罹り、3分の1が
亡くなっている

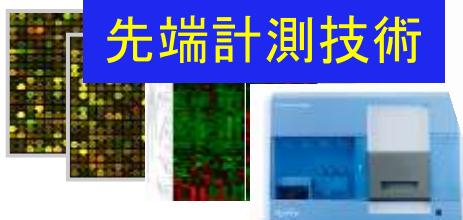


がん研究の範囲は、ゲノム・エピゲノムから
オミックスデータ全体に広がった。さらに、
個人ゲノム時代の到来が目前



$10^3 \sim 10^4$ オーダーの

**大規模臨床サンプル
個人ゲノム**



網羅的オミックスデータ

ゲノム、エピゲノム、遺伝子発現、マイクロ
RNA、プロテオーム、メタボローム解析など
の

大量データ



**走って
なんか
行ける
所じゃ
ない！**



歳をとらない人 その後ろで歳をとっていった人



声の出演
高橋和枝
(2代目:28年間)
1970年1月11日~1998年5月10日

磯野カツオ(11歳、小学5年生)
(テレビ番組「サザエさん」より。1969年10月放送開始)

つまり、

『年寄りはつらいよ』

骨髄異形成症候群で、1999年3月23日
東大医学部附属病院で亡くなる(70歳)。

➤ どんな病気?

- 骨髄で正常な血液を作れなくなる、
代表的な「血液がん」。
- 急性骨髓性白血病への移行が特徴。
- お年寄りに多い。日本で、推定で毎年新たに5千人が罹患。
- **その原因は不明だった。**

➤ 治療法は?

- 骨髄移植しかない。
- しかし、骨髄移植の適応は一般に60歳まで。

みんな歳をとる

60歳、65歳を過ぎても、なんとか働き続け、
互いに支えていかねばならない日本社会が到来する。
そう簡単にがんで寝付くわけにはいかない！



読売新聞 2012年1月30日(月)より

さらに、がんは複雑

- 原発から浸潤、転移へ、がん細胞は変幻。
- がん幹細胞や、正常細胞である血管内皮細胞、免疫炎症細胞などと「つるみ」ながら、
- 薬剤耐性を獲得し、
- 時空間で進化するヘテロな細胞集団
- 複数の原発が進化することもある！



さらに、がんは複雑



気持ちは、ムンクの『叫び』

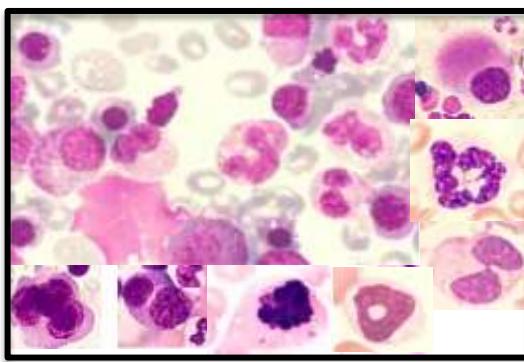
この複雑さをどのように理解すればよいのか？



システムがん

骨髓異形成症候群(MDS)の 原因遺伝子の発見

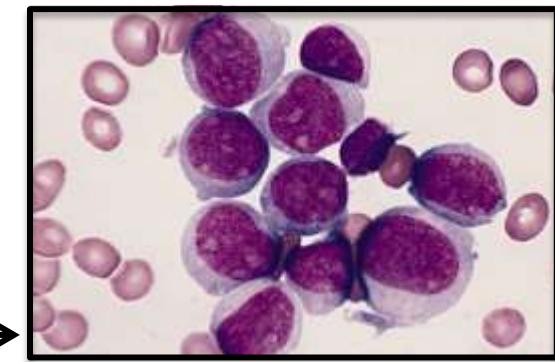
難治性血液がんの克服に向けて



骨髓異形成症候群

A02 小川誠司
(東大医学部附属病院)

→
移行



急性骨髓性白血病

◇いろいろな血液細胞が作
れなくなることで
赤血球減少 → 貧血、動悸、息切れ
白血球減少 → 感染しやすくなる
血小板減少 → 出血しやすい

【共同研究】

A01 宮野 悟(東大医科学研究所)

MDS: myelodysplastic syndromes

システムがん経費で
アップグレード



2010年

Illumina Hiseq2000

スパコンと次世代シーケンサーを使い、がんの新たな原因機構を発見した我が国初の大規模ゲノムシーケンス研究

MDSの原因は不明

次世代シーケンサー Hiseq2000を用いて、29 検体の全エクソン・シー クエンスを行う

ヒゲノム解析センタース パコン上に、超高速、 効率的に変異遺伝子 の候補を検出する耐 故障性データ解析パ イプラインを構築

候補変異を高速リ シーケンス技術を 用いて、多数検体 について効率よく スクリーニングす る技術を開発

高精度変異同定 アルゴリズムを開 発。 血液腫瘍試料に は正常細胞が混 入しているとい う 困難を克服。

- 582例の検体を用いた変異解析で再現確認。
- 変異体を導入HeLa細胞でRNAスプライシング異常が起 こることを確認。
- 変異したスプライシング因子が、造血機能に及ぼす効 果をマウスに骨髄移植して検討。

268個の体細胞 変異を同定

RNAスプライシング 経路の4つの 遺伝子に異常 が頻出

がん研究の歴 史で初めての 知見

原因を解明

2011年9月

データ解析パイプラインを公開



6000コア(75TFLOPS)
1PB高速ストレージ



エクソーム解析・変異解析



融合遺伝子解析





がん研究の歴史に刻まれる発見



TCGA: 2005年に開始。これまでに毎年約50M\$の研究費



Keynote Speechで紹介(2011年11月17日)
Eric Lander, Director of Broad Institute of MIT
and Harvard
Short Biography
BS: Mathematics
MS: MBA
PhD: Molecular Biology

Insights from Genomic Approaches

Functional classes	
Protein and lipid kinases	BRAF (melanoma), PIK3CA (breast), PIK3R1 (GBM), EGFR (lung), FGFR2 (endometrial)
JAK2 (myeloproliferative)	
Lineage survival genes	MITF (melanoma), NKX2-1 (lung), SOX9 (colon)
	SOX2 (squamous lung and esophageal),
Epigenetic regulators	DNMT3A (AML), EZH2 (DLBCL), UTX (many), MLL2/3
Metabolic enzymes	IDH1/2 (GBM and others)
RNA splicing factors	SRSF2, SF3B1, ZRSR2, U2AF1 (MDS), SF3B1 (sideroblastic anemia, CLL)
Translocations	ERG (prostate), ALK (lung)
Surprises	
Notch	Oncogene (T-ALL) and tumor suppressor (squamous)



nature
International weekly journal of science
Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video | Full text
Archive | Volume 478 | Issue 7387 | Articles | Article
NATURE | ARTICLE | previous article | next article |
Frequent pathway mutations of splicing machinery in myelodysplasia
Kenichi Yoshida, Masashi Saneda, Yuichi Shiraishi, Daniel Nowak, Yasunobu Nagata, Ryo Yamamoto, Yusuke Sato, Aiko Sato-Otsubo, Ayana Kon, Masao Nagasaki, George Chalkiadis, Yutaka Suzuki, Masashi Shioseka, Ryoichiro Kawahata, Tomoyuki Yamaguchi, Makoto Ohsu, Naoshi Obara, Mamiko Sakata-Yanagimoto, Ken Ishiyama, Hiraku Mori, Piorian Nolte, Wolf-Karsten Hofmann, Shuichi Miyewaki, Sumio Sugano, Claudia Heferlach, H. Phillip Koefeler, Lee-Yung Shih, Torsten Heferlach, Shigeru Chiba, Hiromitsu Nakachi, Satoru Miyano & Seishi Ogawa
Show fewer authors
Affiliations | Contributions | Corresponding author

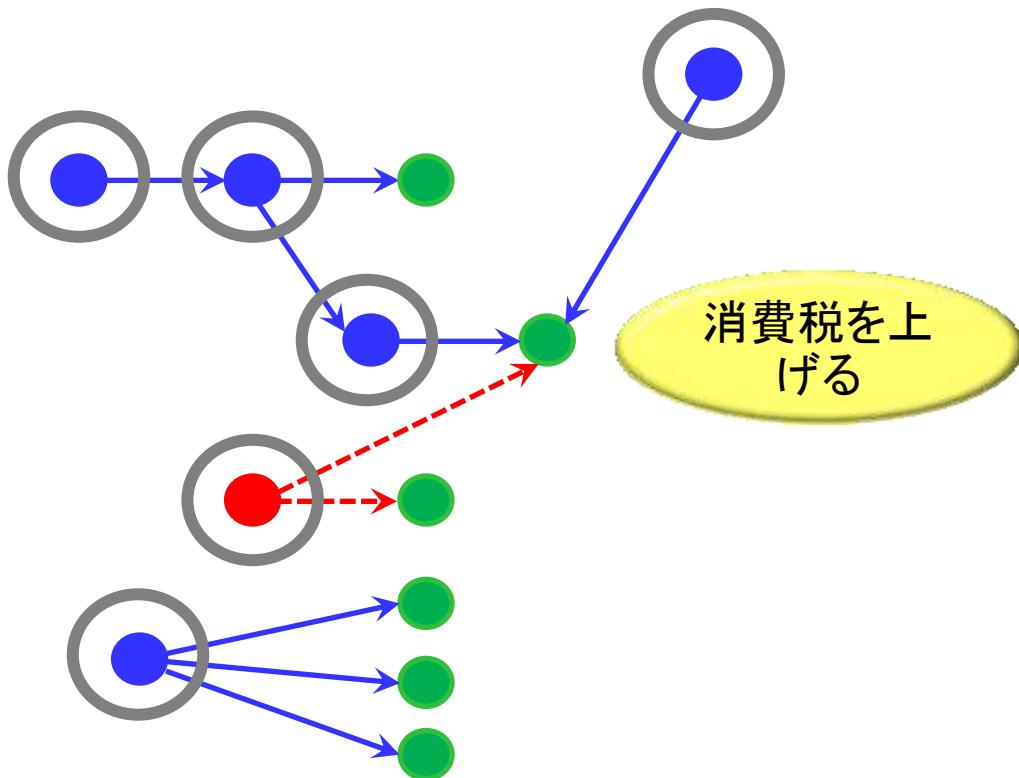
Nature 478, 64–68 (06 October 2011) | doi:10.1038/nature10496
Received: 07 June 2011 | Accepted: 24 August 2011 | Published online: 11 September 2011

この発見は、世界で初めて、骨髓異形成症候群(MDS)の原因遺伝子を発見したという意義だけでなく、「RNAスプライシング」の異常が、がんの発症に関わることを示した世界で初めてのもの

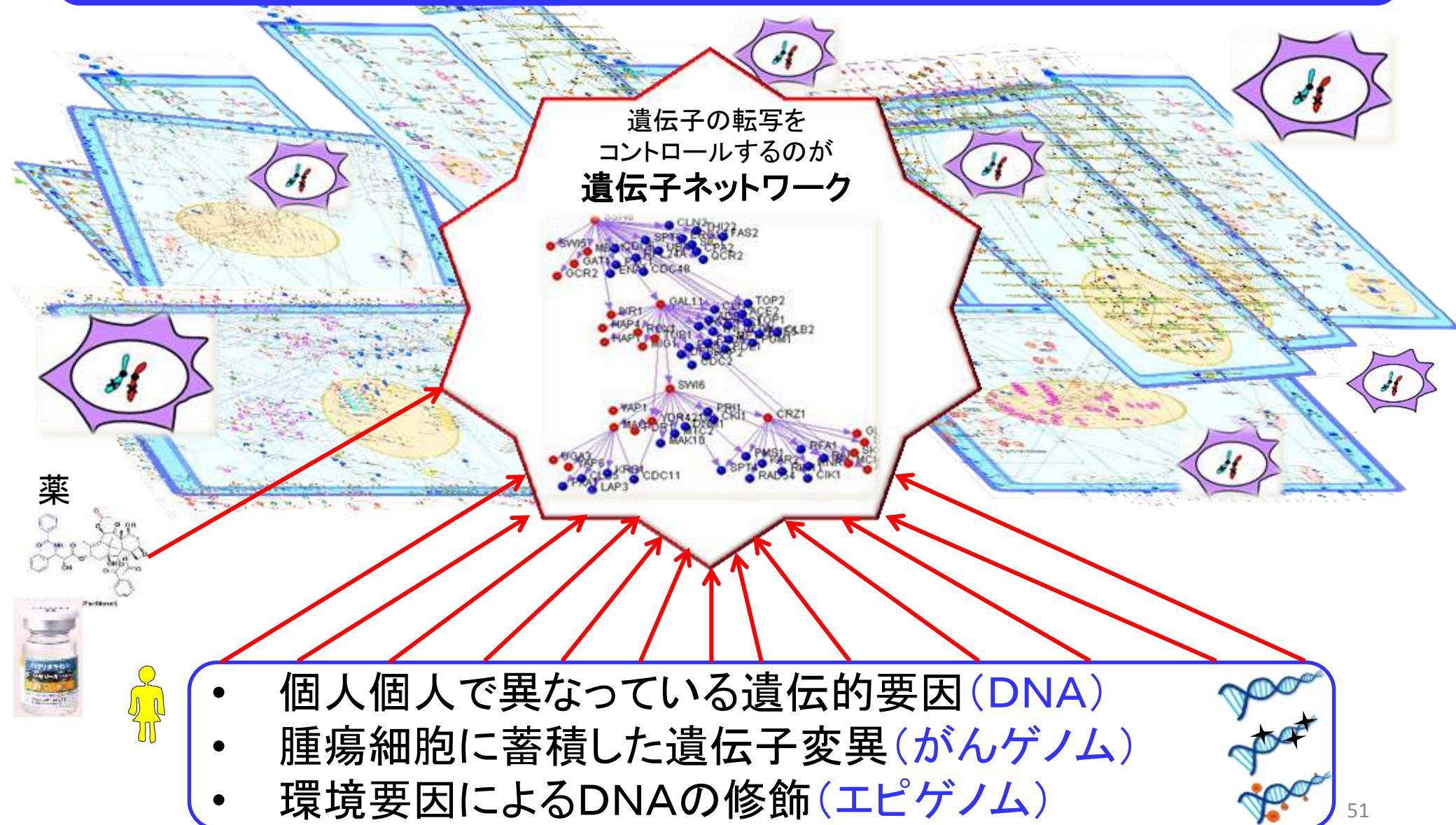
遺伝子ネットワーク

～「職場・社会における人間関係」のようなもの～

- 「働け」と命令する人
- 仕事のじやまをする人



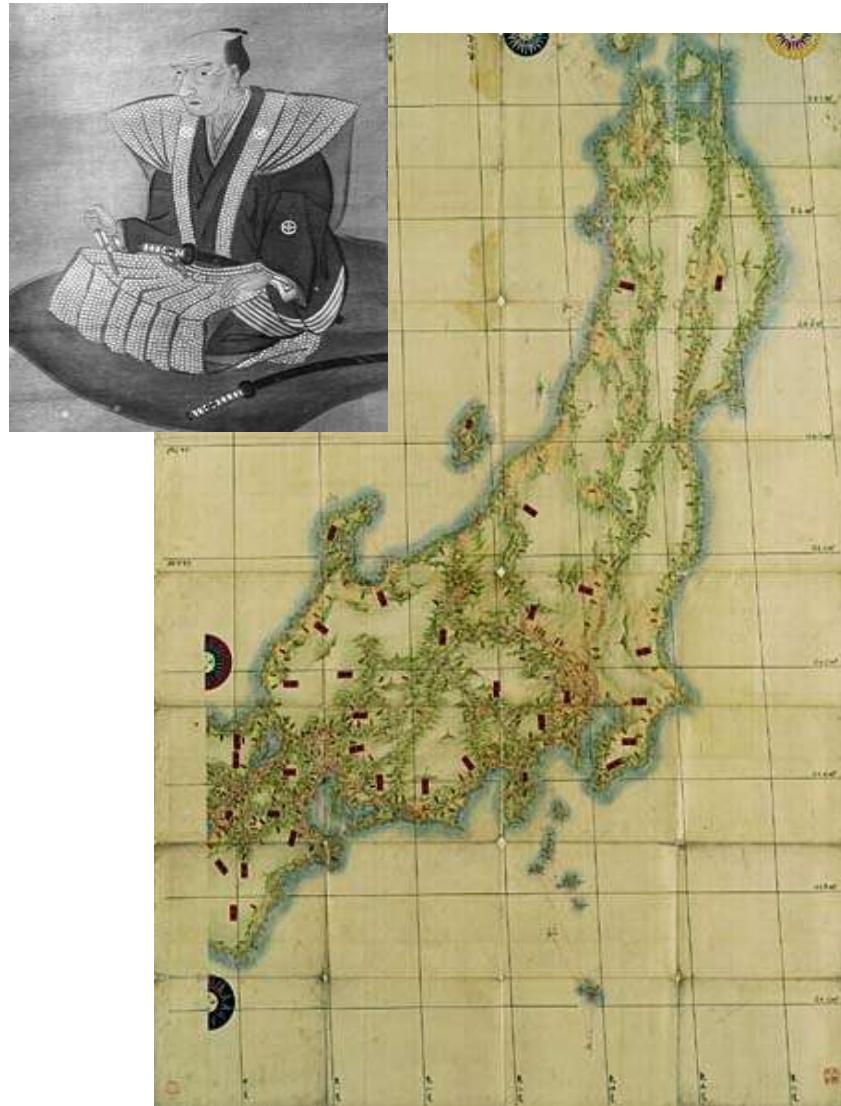
がんの悪性度や治療応答性、副作用の出やすさなどを規定している。



四千万歩の男・伊能忠敬

- 伊能忠敬は、江戸幕府の事業として、1800年から1816年にかけて全国を歩いて測量をし、1821年に「大日本沿海輿地全図」が幕府に納められたといいます。伊能忠敬はその完成を見ずに1818年に死去ましたが、その後、仕上げの編纂作業が行われ、全部で21年の歳月をかけてこの地図は完成しています。
- それから200年後の現在、Google™マップに象徴されるように、だれもが世界中をナビゲーションして、目的の場所に行くことができます。

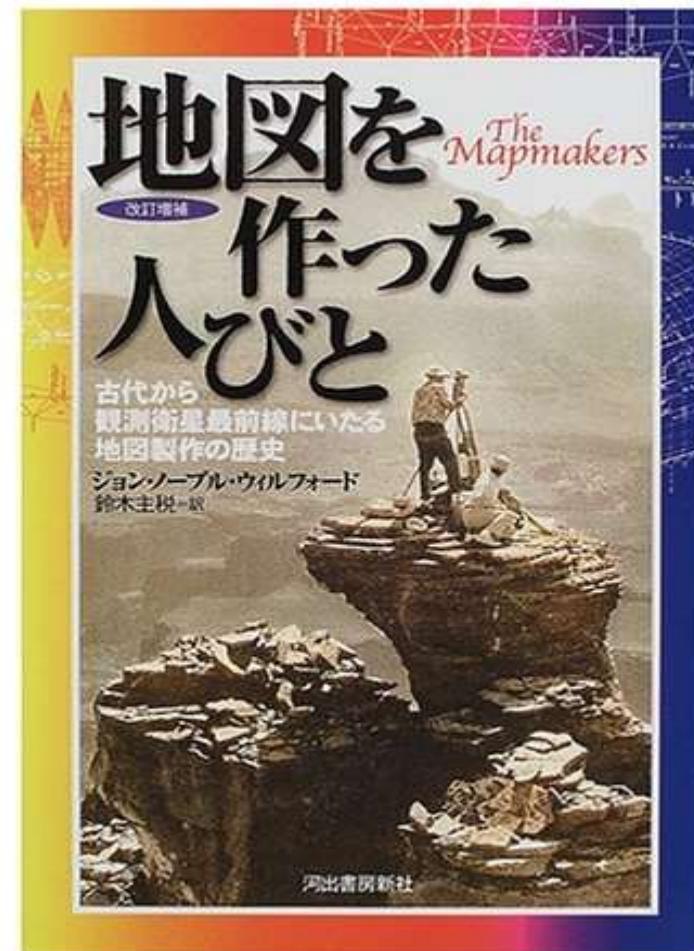
2001年のNHKの正月時代劇「四千万歩の男・伊能忠敬」
(原作:井上ひさし「四千万歩の男」(講談社))として放映



地図を作った人びと

－古代から観測衛星最前線にいたる地図製作の歴史－

- John Noble Wilford著
- ピューリツサー賞受賞作



「がんの遺伝子ネットワークアトラス」 を描くためのデータ

- 全遺伝子の発現状態をmRNAなどの転写産物量としてDNAチップなどで計測したものです。マイクロRNAも含まれています。先ほどのシーケンサーによっても定量計測できるようになりました。
- 先ほどのたとえで言いますと、**それぞれの人の働きぶりを数値化したものです。**
- 遺伝子ネットワークを推定するアルゴリズムは、**職場・社会における人の働きぶりデータから、人間の制御構造とその影響力を推定します。**

スパコンがあぶり出した再発リスクに関わる肺がん患者のパーソナル遺伝子ネットワーク

私の肺がんのシステムはどうなっているのか？
再発するの？

データ: 226症例の肺線がん患者の遺伝子発現データ

(国立がん研究センター横田淳先生、河野隆志先生との共同研究)

目的: 再発リスクをモジュレーターとしたときの 226 症例各々の遺伝子ネットワークを推定し、再発リスクが高い患者と低い患者のシステムの違いを比較



数百人のがん検体



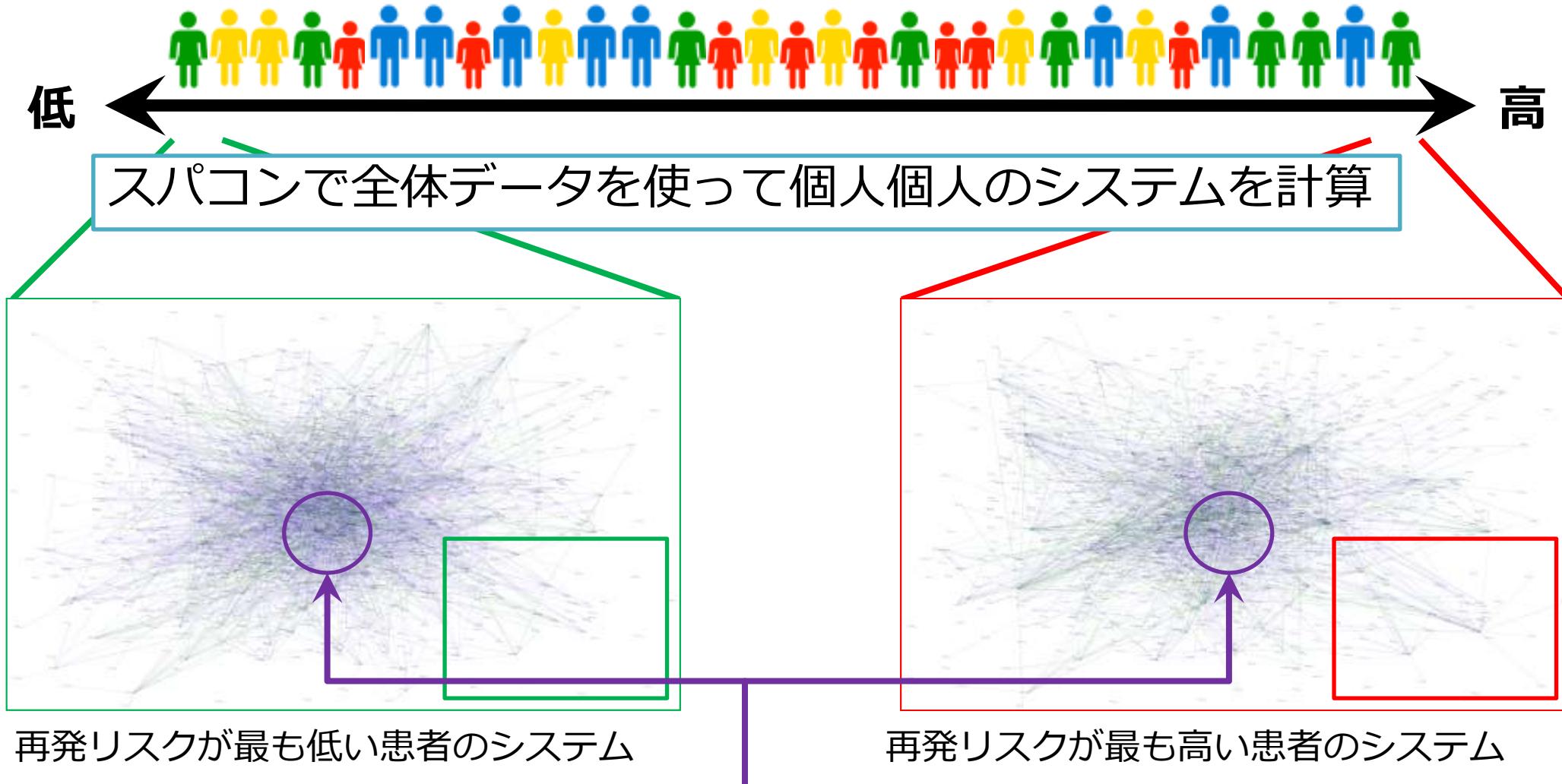
低



高

再発リスク

再発リスク



CTGF (Connective Tissue Growth Factor)

増殖・分化を制御するTGF- β の下流因子で、TGF- β の間質纖維化促進作用を仲介
増殖促進、遊走、細胞外基質算出、血管新生作用を呈する

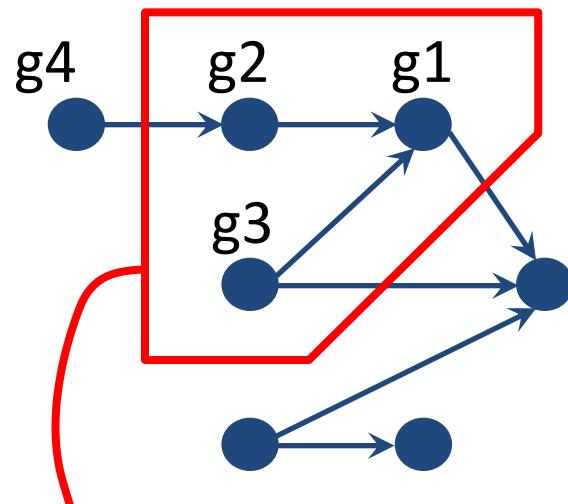
“因果応報”

Bayesian Network+予後解析

- たくさんのサンプル(たくさんの人の活動スナップショット)から1つのネットワークを作る
「遺伝子間の因果の関係図」
- 予後データ(それぞれの人生の結果)から因果の図を解析すると
「応報」の原因がみえてくる。

“仕事場における人間関係のようなもの”

Bayesian Network

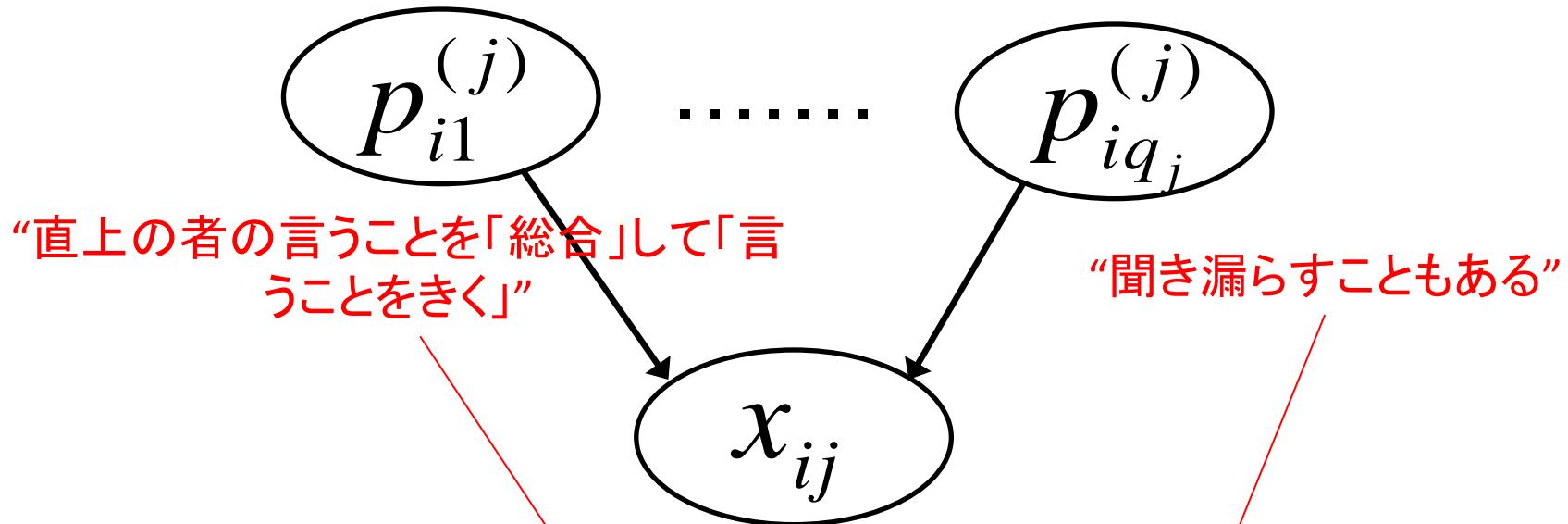


- Markov assumption
「直上の言うことしか耳を傾けない」
- ネットワーク構造がわかっていれば、データがたくさん観測されると、
「言うことをきく」確率がわかる。

$$x_{i1} \iff \mathbf{p}_{i1} = (x_{i2}, x_{i3})^T$$

$$f(x_{i1}, \dots, x_{ip} \mid \boldsymbol{\theta}_G) = \prod_{j=1}^p f_j(x_{ij} \mid \mathbf{p}_{ij}, \boldsymbol{\theta}_j)$$

Nonparametric Regression



We consider the additive regression model:

$$x_{ij} = m_1(p_{i1}^{(j)}) + \dots + m_{q_j}(p_{iq_j}^{(j)}) + \varepsilon_j,$$

where $\varepsilon_j \sim N(0, \sigma_j^2)$ and $\mathbf{p}_{ij} = (p_{i1}^{(j)}, \dots, p_{iq_j}^{(j)})$.

Here $m_k(\cdot)$ is a smooth function from R to R.

“数学的にはこんな感じになります”

Nonlinear Bayesian network model

$$f(x_{i1}, \dots, x_{ip}; \boldsymbol{\theta}_G) = \prod_{j=1}^p f_j(x_{ij} | \mathbf{p}_{ij}; \boldsymbol{\theta}_j),$$

$$f_j(x_{ij} | \mathbf{p}_{ij}; \boldsymbol{\theta}_j) = \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left\{-\frac{(x_{ij} - \mu_{ij})^2}{2\sigma_j^2}\right\}$$

$$\mu_{ij} = m_1(p_{i1}^{(j)}) + \dots + m_{q_j}(p_{iq_j}^{(j)})$$

$$= \sum_{k=1}^{q_j} \sum_{m=1}^{M_{jk}} \gamma_{mk} b_{mk}^{(j)}(p_{ik}^{(j)})$$

たくさん
のパラメータ

“難しいのはみんなの仕事ぶりデータからネットワーク構造を推定すること”

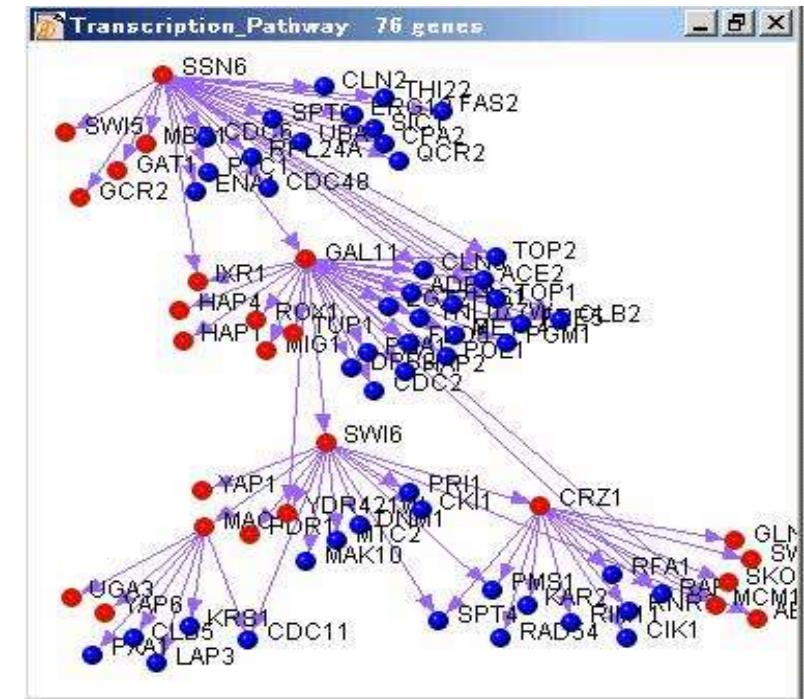
Criterion for selecting good networks

BNRC Score
Bayesian Network and Nonparametric Regression Criterion

$$\begin{aligned} \text{BNRC}(G) &= -2 \log \pi_G \int \prod_{i=1}^n f(\mathbf{x}_i; \boldsymbol{\theta}_G) \pi(\boldsymbol{\theta}_G | \lambda) d\boldsymbol{\theta}_G \\ &= -2 \log \pi_G - r \log(2\pi n^{-1}) \\ &\quad + \log|J_\lambda(\hat{\boldsymbol{\theta}}_G)| - 2nl_\lambda(\hat{\boldsymbol{\theta}}_G | \mathbf{X}_n) \end{aligned}$$

この“BNRC score”が小さいネットワーク構造と
パラメータを探索する。 “スパコンがここで必要になる”

Bayesian Network + Nonparametric Regression



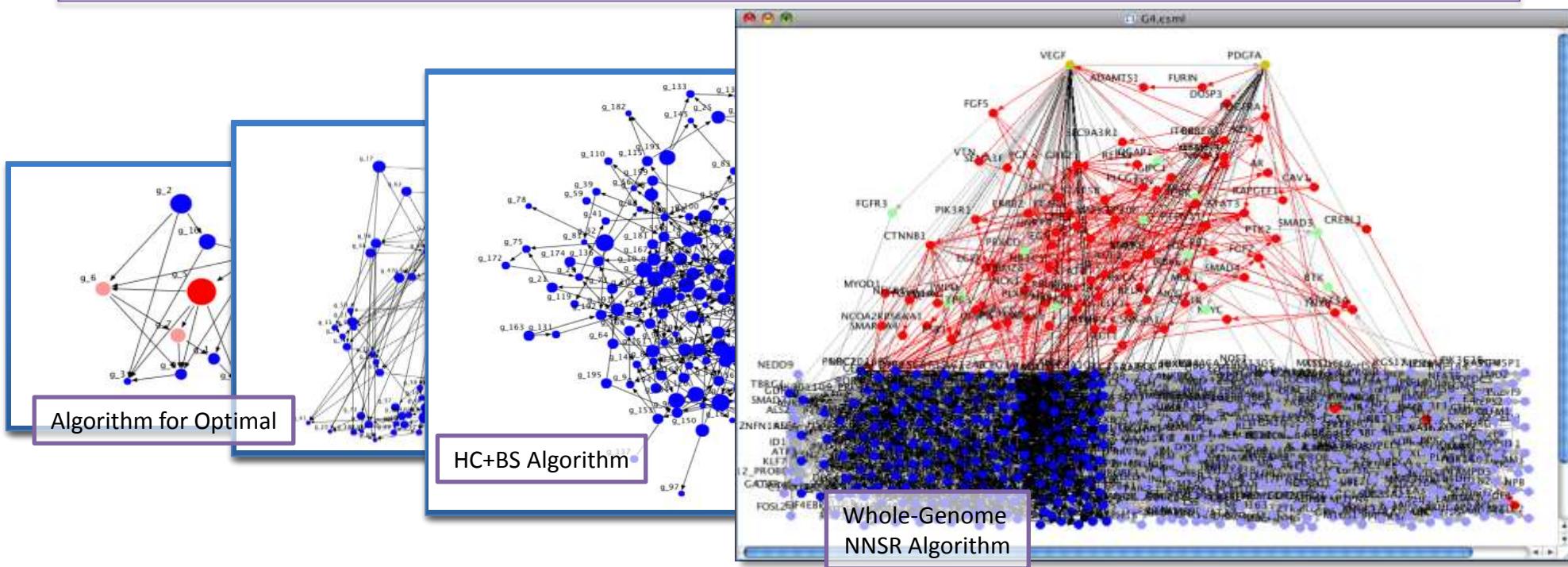
遺伝子ネットワーク

1. Imoto, S., Goto, T., Miyano, S. Estimation of genetic networks and functional structures between genes by using Bayesian network and nonparametric regression. Pacific Symposium on Biocomputing. 7:175-186, 2002.
2. Imoto, Kim, Goto, Aburatani, Tashiro, Kuhara, Miyano (2003). Bayesian network and nonparametric heteroscedastic regression for nonlinear modeling of genetic network. *J. Bioinformatics and Comp. Biol.*, 1(2), 231-252

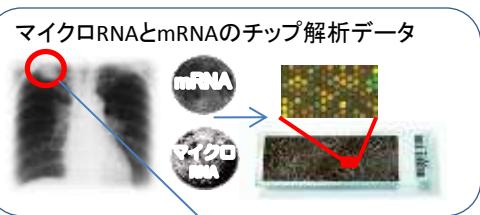
A Series of Programs on Supercomputer for mining gene networks of size from 30 to 20,000 (genome-wide)

Optimal to Locally Optimal

WR: Optimal Bayesian Networks of 32 Nodes
April 2010



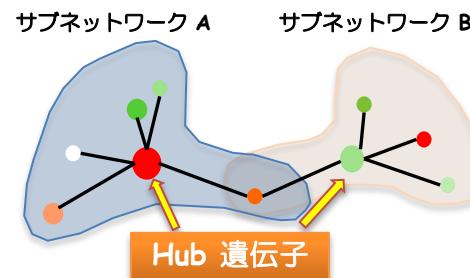
肺腺がんの予後の良・不良のスイッチを切り替えているハブ遺伝子とサブネットワークの発見 NKX2-1の実験による検証と新規マイクロRNA miR-30cとその標的遺伝子RRM2の発見 (SiGN-BNの応用) 研究成果は発表準備中



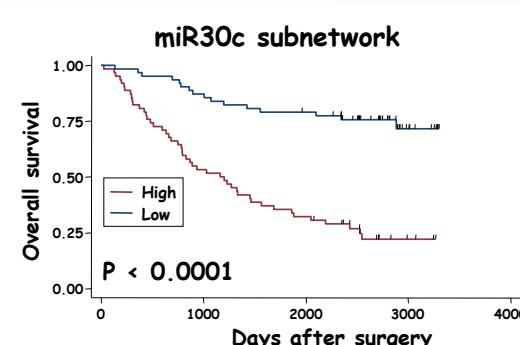
非線形回帰ベイジ
アンネットワーク

SiGN-BN

肺がんのマイクロRNA遺伝子ネットワーク



予後データ



ハブ機能の実験検証

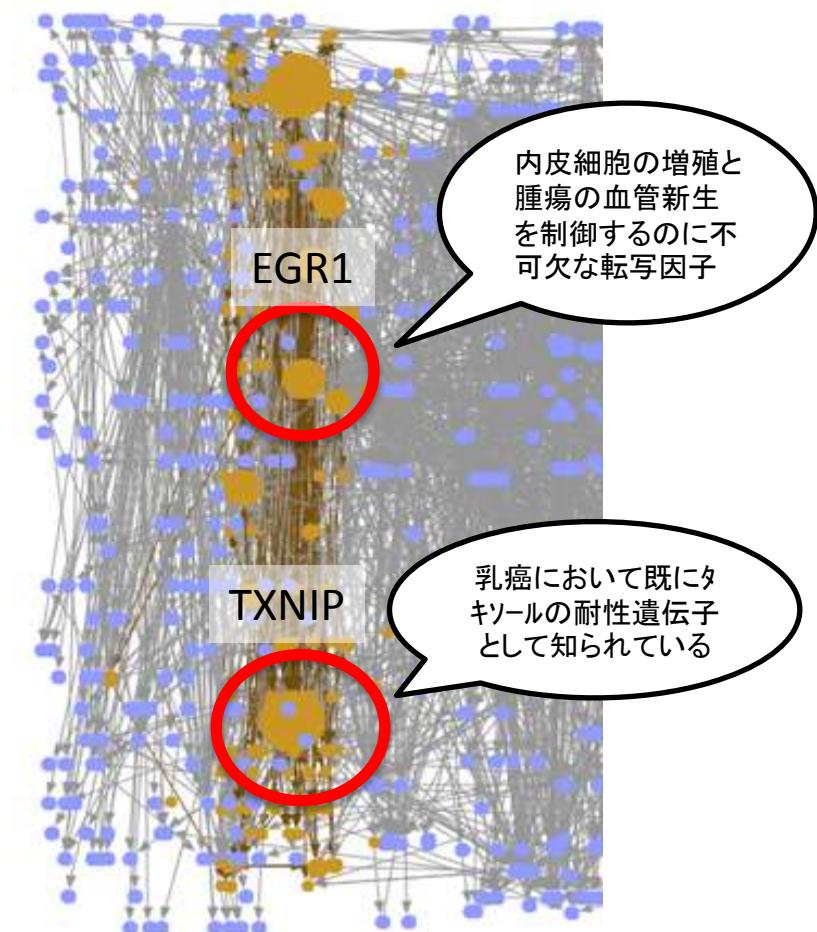
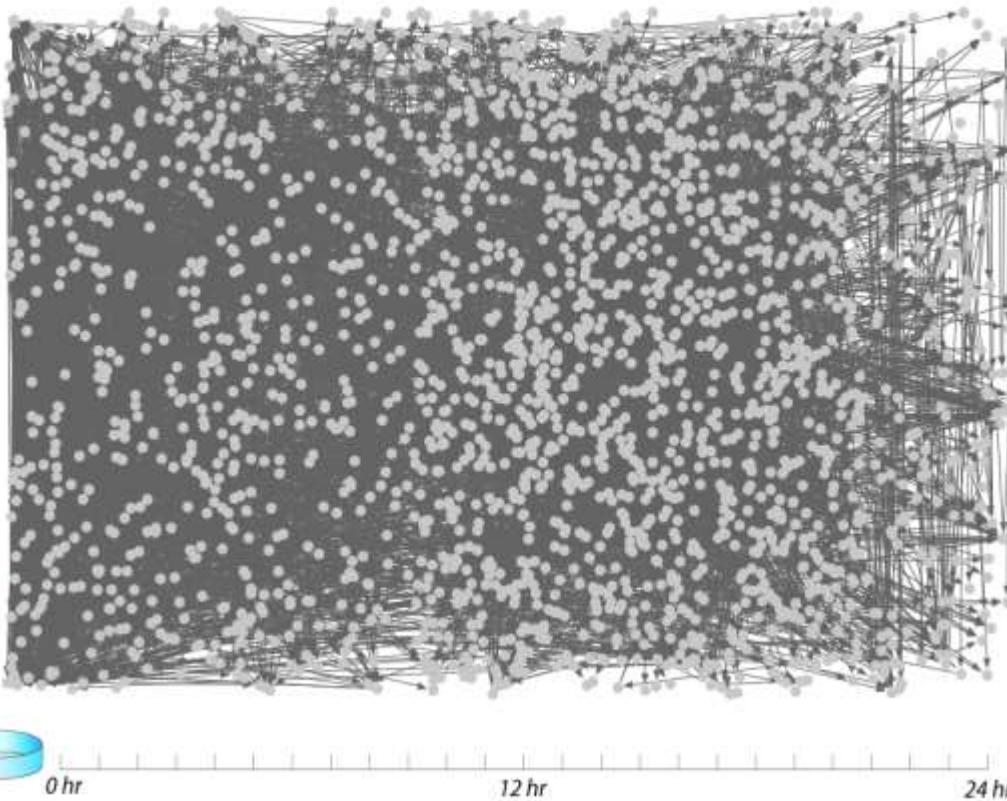
再発・死亡と優位に関連する14個のサブネットワークと、そのハブ遺伝子が見つかる

予後予測力プラン・マイヤーダイヤグラムのp値が小さくなるサブネットワークを網羅的に探索

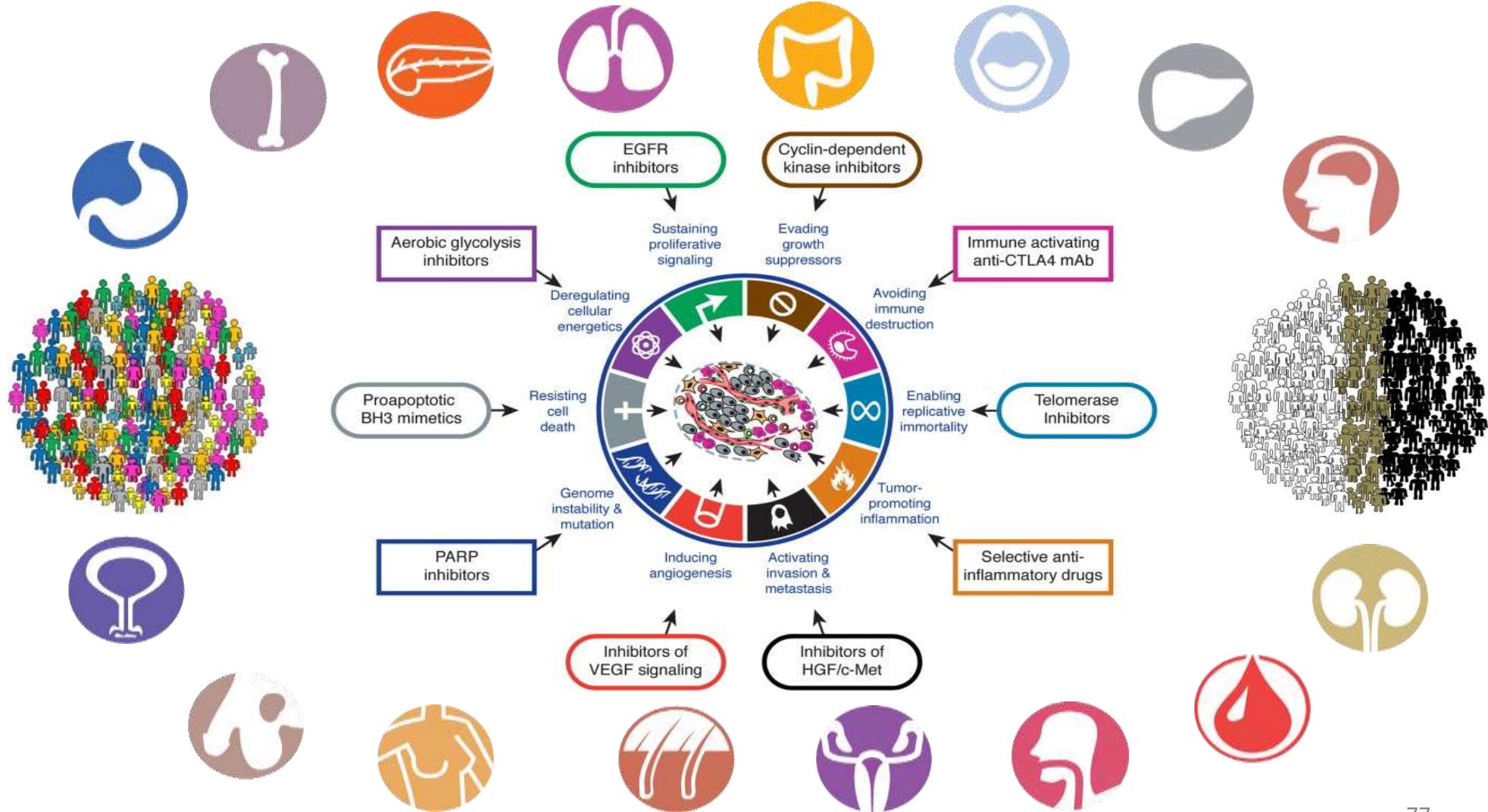


高橋隆教授(名大医)と連携し、SiGN-BNで解析

スパコンが描出した抗がん剤応答 がん遺伝子ネットワーク



がん遺伝子ネットワークからなる 『The Cancer Network Galaxy』を京で構築したい 「京」を持ってる日本しかできない発想…



世界初の大規模がん遺伝子ネットワークデータベースがスタート (SiGN-BNで開始)

コンテンツ: 京を約30ノード年使い計算した8000遺伝子からなる256の様々ながんの遺伝子ネットワーク



<http://tcng.hgc.jp/>

TCNG The Cancer Network Galaxy^{0.13}

Database of Cancer Gene Networks from Public Gene Expression Data

[Home](#) [About](#) [Browse](#) [Help](#) [Download](#) [Publications](#)

[Contact](#)

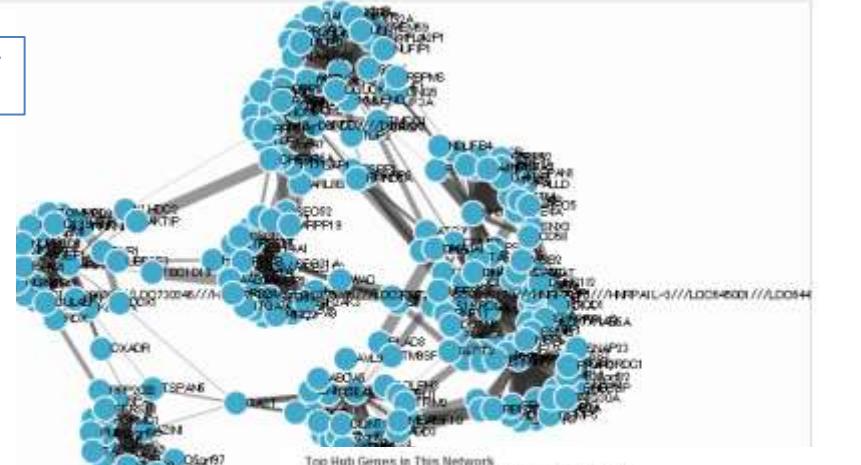
search query

search

Gene Network

GSE14095_top8000 - GSE14095 - SiGN-BN NNSR

Network Name	<i>GSE14095_top8000 - GSE14095 - SiGN-BN NNSR</i>
Array Set	<i>GSE14095 (See on GEO)</i>
Array Set Source	<i>GEO</i>
Gene Set	<i>GSE14095_top8000</i>
Gene Set Description	<i>Gene expression signature and the prediction of liver metastasis in colorectal cancer by DNA microarray</i>
Number of Nodes	<i>8000</i>
Network Method	<i>SiGN-BN NNSR</i>
Downloadable Files	<i>text format - GSE14095.txt CSML format - GSE14095.csml NODELIST format - GSE14095.nl.txt</i>

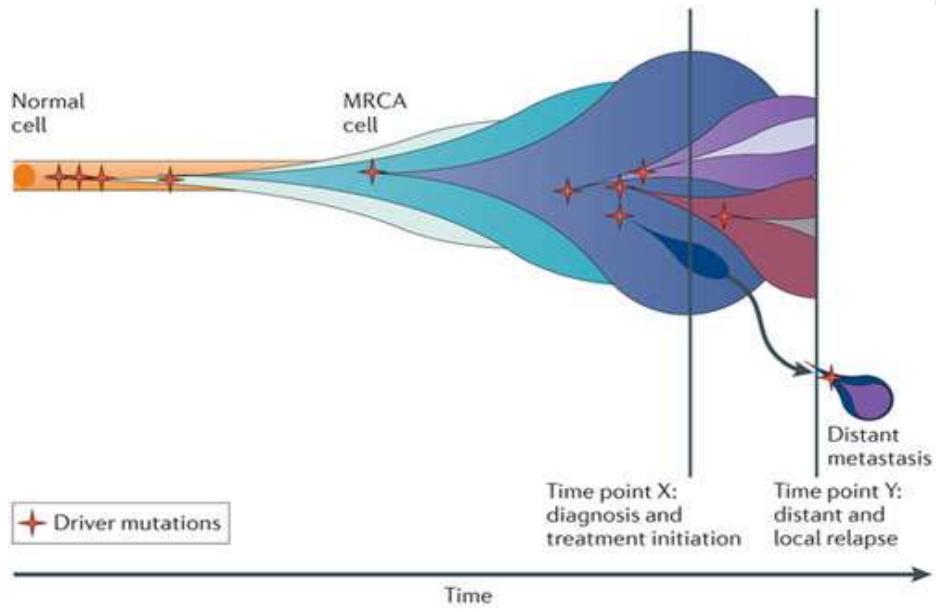


Top Hub Genes in This Network
(Click on a gene name to see the network around the gene.)

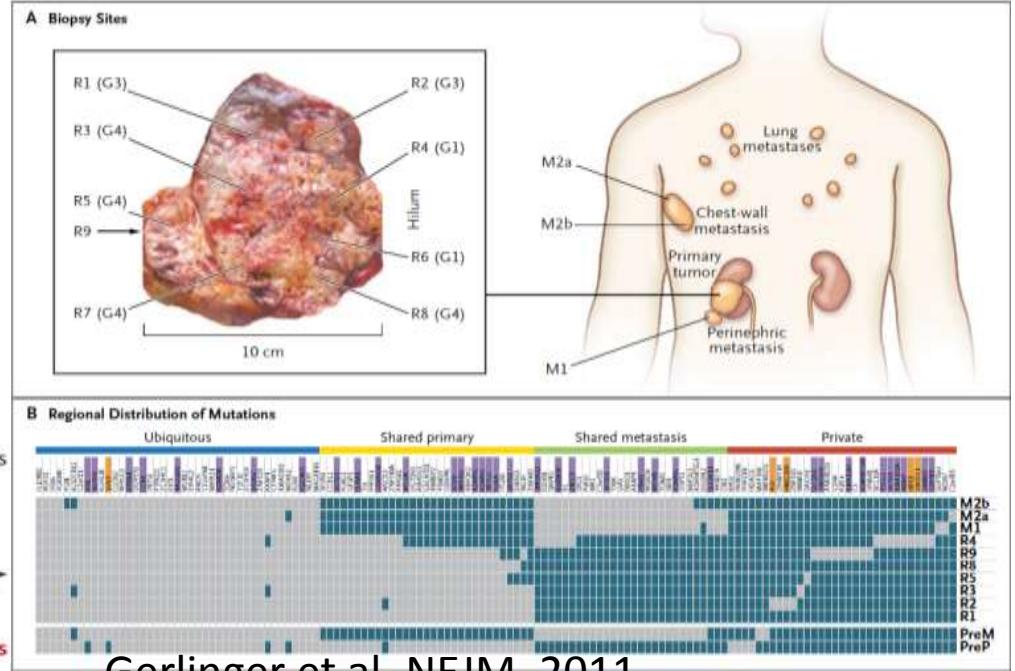
Hub Rank	Node	Children	Parents	Degree	Betweenness
#1	HCK	155	4	159	14657.701
#2	EGR1	104	4	108	611713.013
#3	CDH1	85	6	91	101758.013
#4	TMEM108B	84	4	88	30654.053
#5	EXOSC3	84	5	87	100148.881
#6	TMEM108B	84	5	89	27052.843
#7	RAE1	79	6	87	180033.052
#8	CDH1	77	4	81	937040.114
#9	SIT	73	5	78	263278.024
#10	LYRG6D	71	6	77	114379.159
#11	MED27	70	3	73	68860.038
#12	AGTR2	68	3	71	123788.404
#13	C1orf6	67	5	72	199972.383
#14	DCUMLD1	67	6	73	527079.669
#15	RAB7B	65	3	68	442260.943
#16	RAB31	65	2	67	4039.98
#17	REG4	63	3	66	182187.078
#18	CANT1	62	3	65	6180258.137
#19	POURSE	62	9	64	5604.003
#20	REMB4	61	4	65	91385.168
#21	CARS	60	5	65	161026.429
#22	AP1S3	60	3	63	438374.74
#23	SIMP3	60	5	62	2667.011
#24	TDPBP // TDGF1	59	5	64	80589.318

挑戦的課題は

個々人のがんの病態(システム異常)を大規模スケール(多人数)で網羅的に(ゲノムワイド)に、3次元的、4次元的に解析すること



Nature Reviews | Genetics



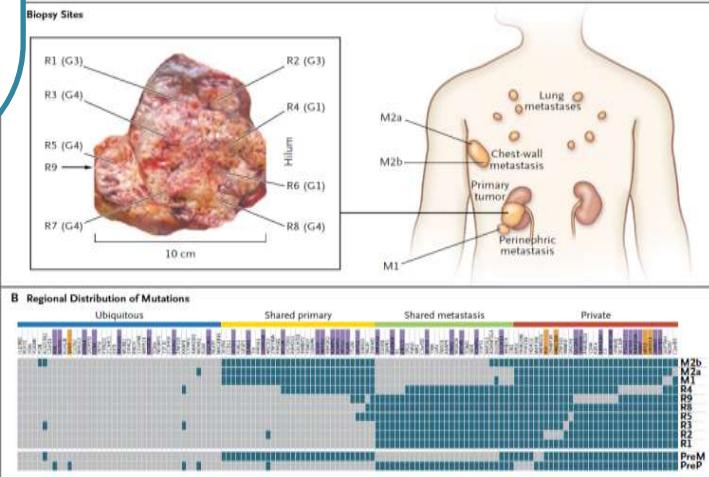
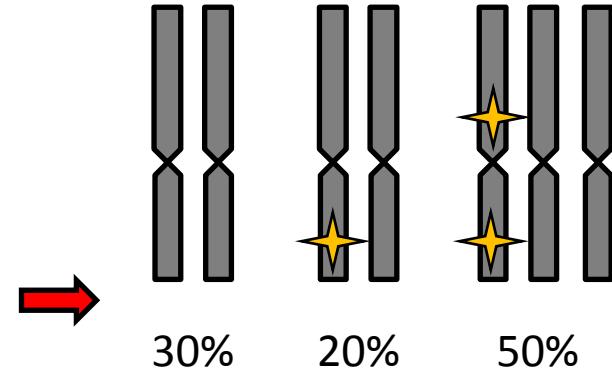
Gerlinger et al., NEJM, 2011

がんの3次元的ヘテロ構造を暴く

@HWUSI-
EAS1765_0005_FC628ALAXX:4:1:3976:1041
TTCACCCATCCACAAGTTCAAAGTTATAGTCTCCAGGG
CAGACANGTN
+
hghhhhhhhhhhhhhhhhhhhhhhhhhhhfffff
hhhhhaB^B
@HWUSI-
EAS1765_0005_FC628ALAXX:4:1:5060:1043
TGTGTCATTTCTGGCTGCCGATCTCACGCATCTCTAAA
GTGAAAATT
+
Hhhhhhhhhhhhhhhhhhhhhhhhhhhgggggggggggg
ghhhghhhh
@HWUSI-
EAS1765_0005_FC628ALAXX:4:1:3976:1041
TTCACCCATCCACAAGTTCAAAGTTATAGTCTCCAGGG
CAGACANGTN
+
hghhhhhhhhhhhhhhhhhhhhhhhhhfffff
hhhhhaB^B
@HWUSI-
EAS1765_0005_FC628ALAXX:4:1:5060:1043
TGTGTCATTTCTGGCTGCCGATCTCACGCATCTCTAAA
GTGAAAATT
+
hhhhhhhhhhhhhhhhhhhhhhhhhhgggggggggggg
ghhhghhhh

数理モデリングとスパコン

- 次世代シークエンサーデータから、がんゲノムの主要な細胞集団、変異群を推定。
- 個人の複数部位及び再発がんシークエンスに対応したモデリングにより、クローンの組成の変化をより3D-4D的にモデリングする。



Gerlinger et al., NEJM, 2011

さらに計算能力があると道が拓けるのだが…

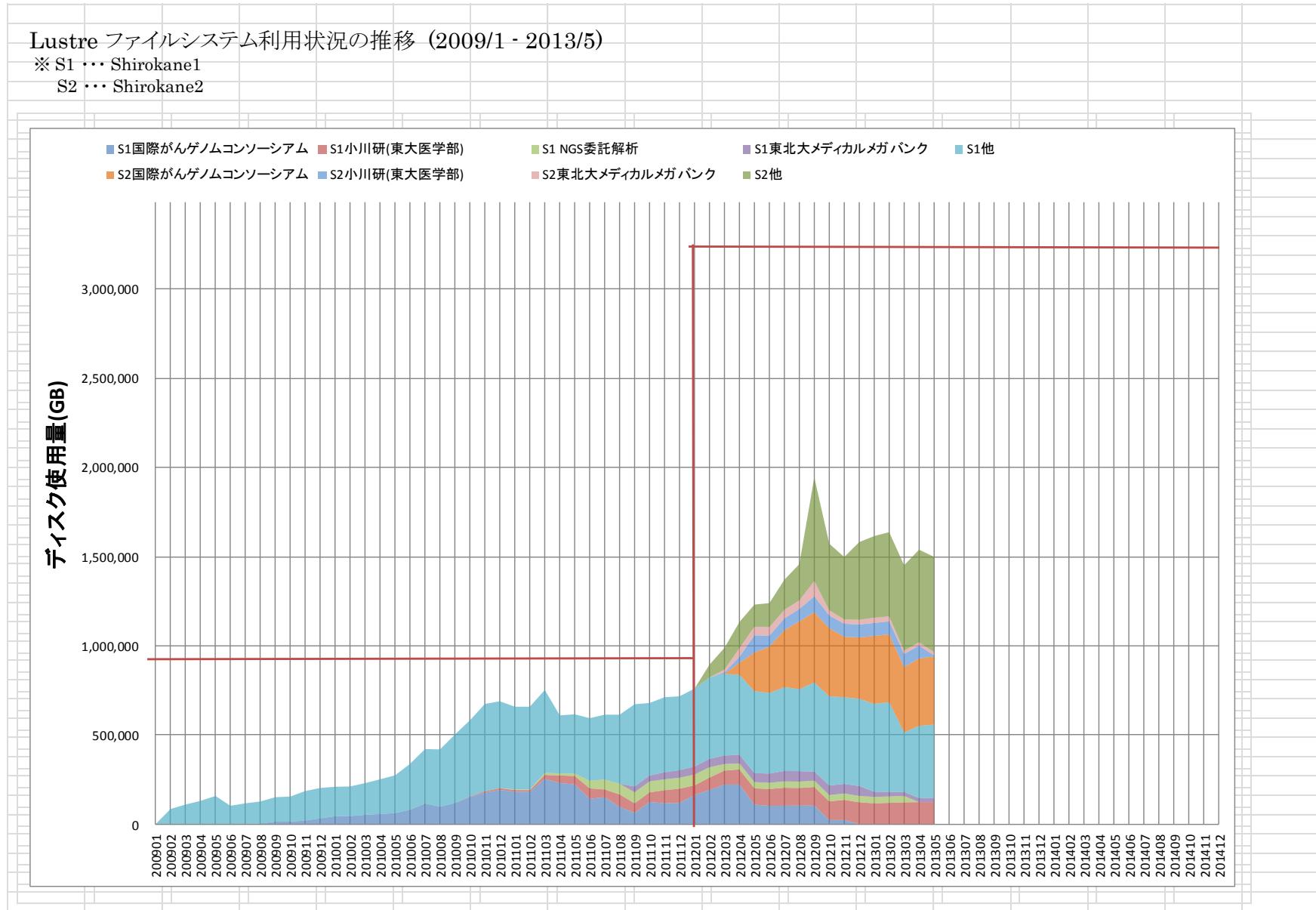
- 40部位検体に対する推定では、1億エクソン・リード
× 40検体 = 40億エクソン
- 40億エクソンを対象とした、がんゲノムの大規模ベイズモデルの構築

⇒

- がんの進化と転移・再発について、がんの3次元的サブクローンヘテロ構造の理解が初めて可能になる。
- 個人のがんの進展、転移、再発のプロセスを捉えることができ、抗がん剤の適用プロトコルなど、治療方針の戦略策定へ可能性を拓く。

迫り来る恐怖

シークエンサーデータによる高速ストレージ量の増化 残り2年弱をどう乗り切るか…5年先は？



さらに
東大医科研財政難のため、2015年
からのスパコン経費30%削減決定

人件費を除いた東大医科研運営費交付金の55%
がヒトゲノム解析センターの経費

東京電力の 電気料金の値上げは深刻(30%)

アベノミックスによる
1ドル80円→120円

は次期システム導入の大きな障害
(Intel, storage, etc.)

追い打ち、消費税10%

ゲノム情報に基づく個別化医療を推進するために必要な計算量と データ量の見積もり(オールジャパン)(①と②の実施)

	2012	2013	2014	2015	2016	2017	2018	2019	2020
ヒト生命データ解析	演算性能	250TFLOPS	1PFLOPS		256PFLOPS		1ExaFLOPS		
	ストレージ	16PB	128PB	256PB	512PB	1ExB	2ExB		
	サンプル数	500	2,000	5,000	10,000	50,000	100,000	300,000	500,000
		初期		中期			後期		
1サンプル当たりのデータ	1TB		1TB～10TB		10TB～100TB				
シークエンサー技術	次世代シークエンサー (HiSeqシリーズなど)	第三世代シークエンサー(PacBioなど:消滅) 第四世代シークエンサー(シリコンシークエンサー) 第X世代シークエンサー				融合技術型超小型シークエンサー			

① 個別化医療推進のための情報基盤の整備

- ヒト大規模ゲノム関連データベース(数十万人)、副作用情報データベース、大規模生命・医療データ解析技術、ソフトウェア等の情報基盤技術を整備。このために、スーパーコンピュータシステム(1エクサ・フロップス)、及び大規模ストレージ(2エクサ・バイト)をインフラとして設備。

② 個別化医療のための次世代ゲノム医学研究

- 超高速シークエンサー技術等を駆使して、個人個人のゲノム情報・エピゲノム・トランスクriptーム・プロテオーム・メタボロームなどの違いと、がんや成人病等の病気、薬、環境因子との繋がりを解明し、それを診断、予防、治療へと翻訳する最先端研究の実施。

ヒトゲノム解析センター時期スパコンの最低限の姿だが…

		2015.1	2016	2017	2018.1	2019	2020.12
ヒト生命データ解析	演算性能	1PFLOPS			2PFLOPS		
	ストレージ	20PB	30PB	50PB	80PB	90PB	100PB
	Lustre 25PB	Near Line 25PB		Lustre 50PB	Near Line 50PB		

リース経費(6年リース) : 120億円
現在(6年リース) : 60億円弱
医科研財務は今後70%に減 : 42億円

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。

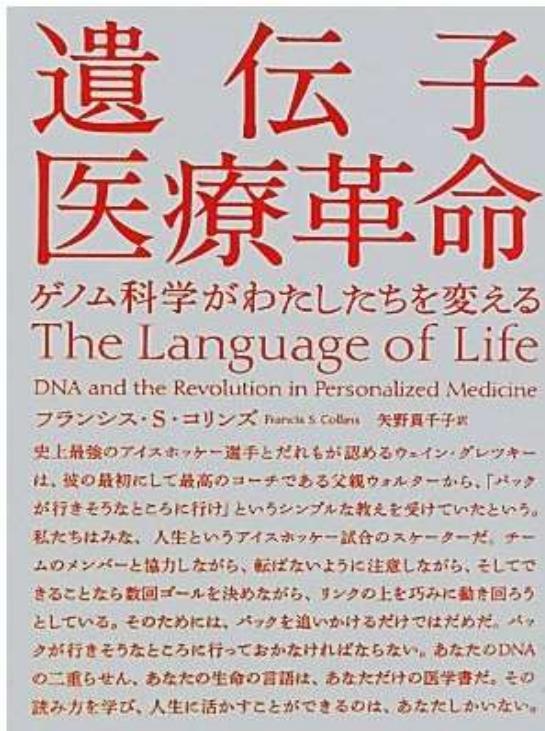
あなたのがんにぴったり合った薬があつて
良かったですね。

- 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。



だれもが保険証を持っているように自分のDNA情報をもつ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



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

