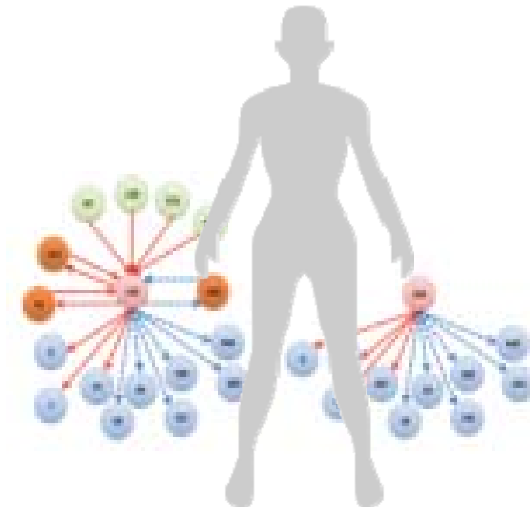
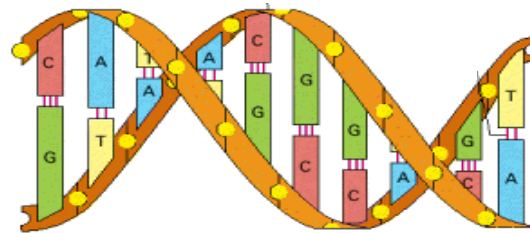
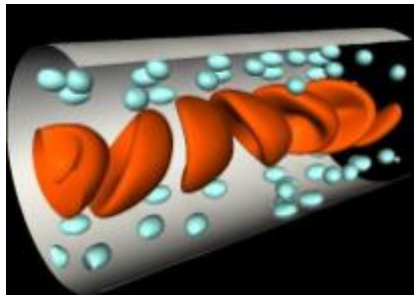


スーパーコンピュータが 解き明かす生命の不思議



秋山 泰 (Yutaka Akiyama)

東京工業大学 大学院情報理工学研究科 教授



「京」 K computer （理研 計算科学研究機構 神戸）

8コアCPU × 88,128ノード

毎秒当たり“1.051京回”の計算（2011年11月 世界一）

「^{けい}京」と「10ペタ」

- 壹 1
- 万 10^4
- 億 10^8
- 兆 10^{12}
- 京 10^{16} 2011年
- 垓 10^{20}
- ■ 10^{24}

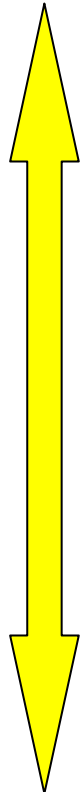
漢字文化圏では1万倍ごとの単位

- One 1
- Kilo 10^3
- Mega 10^6
- Giga 10^9 1983年
- Tera 10^{12} 1997年
- Peta 10^{15} 2008年
- Exa 10^{18}
- Zetta 10^{21}
- Yotta 10^{24}

西洋では1千倍ごとの単位

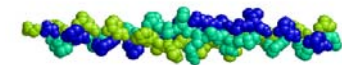
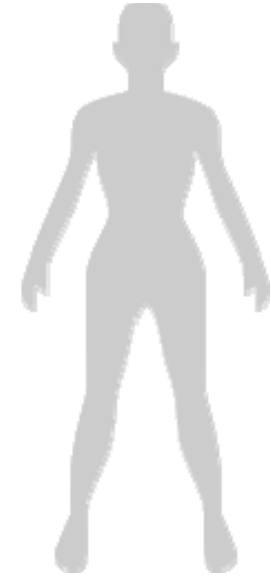
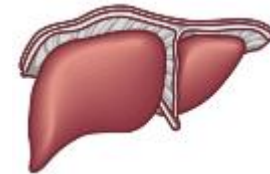
生命は複雑な階層型システム

大

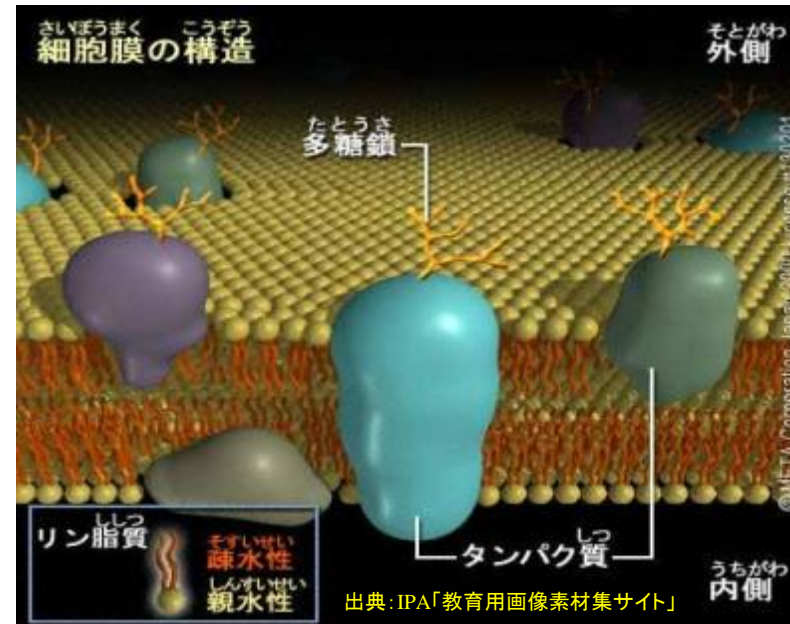
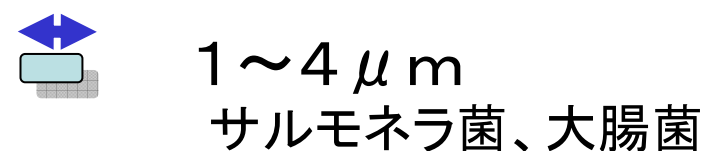
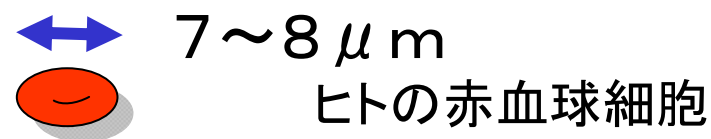
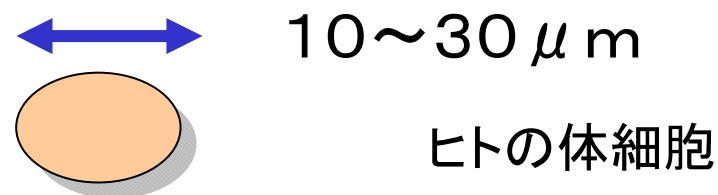
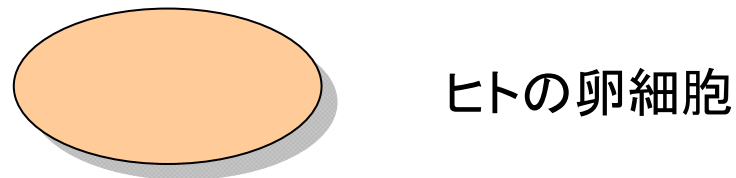
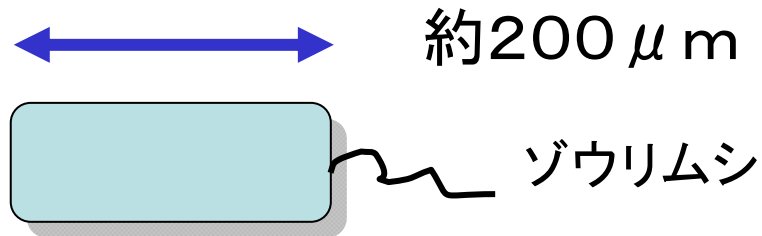


小

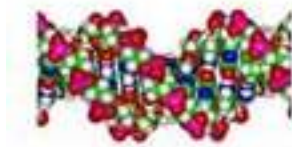
- 生体 (body) ヒトの体には60兆個の細胞
- 器官系 (organ system)
- 器官 (organ)
- 組織 (tissue)
- 細胞 (cell) 生命の基本的単位
- 細胞内小器官 (organelle)
- 分子 (molecule)



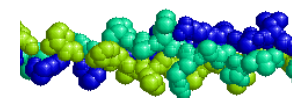
細胞という“微小”システム



細胞膜: 厚み60~80 Å



DNA二重らせん
 直径20 Å
 (ヒト細胞内 全長約2m)

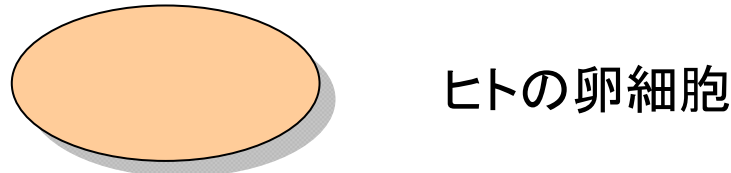
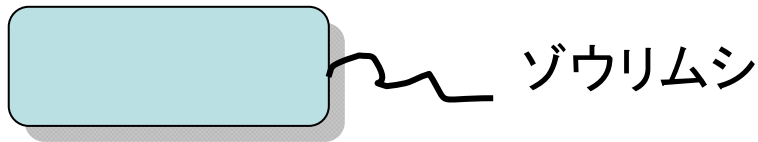


コラーゲン三本鎖
 直径15 Å

細胞という“巨大?” システム

スケールを1億倍してみると、

約20km



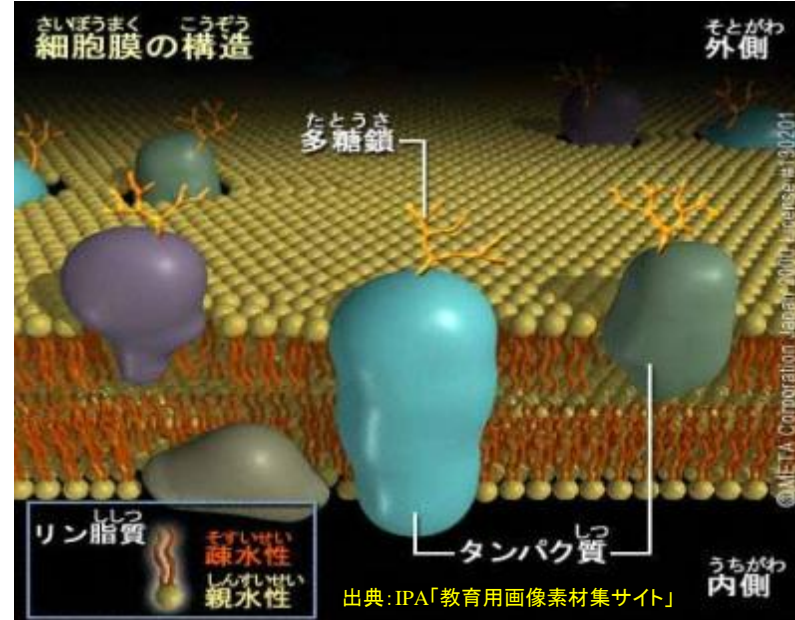
1~3km



700~800m
ヒトの赤血球細胞

100~400m
サルモネラ菌、大腸菌

U.S.S. Enterprise
300 - 600 m



細胞膜: 厚み60~80cm

DNA二重らせん
直径20cm
(ヒト細胞内 全長約20万km)

コラーゲン三本鎖
直径15cm

細胞や人体を理解するという挑戦

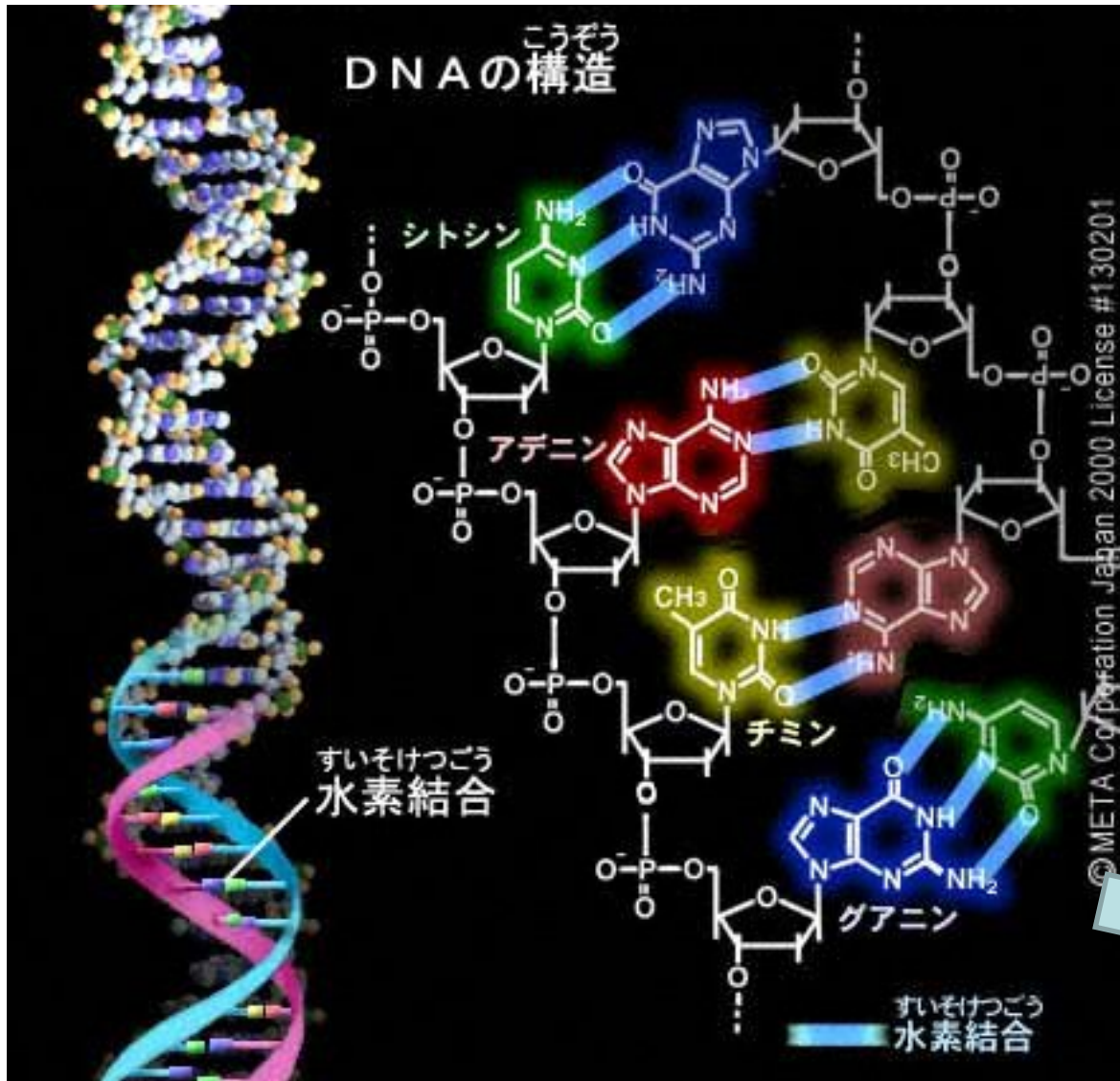
- 部品の尺度から見れば生命は“**超巨大システム**”
- 構成している要素の種類が人工物よりも豊富
ヒト： 約2.5万種類の「遺伝子」(設計図)
約10万種類の「タンパク質」(部品)
- 設計図が少しずつ違う多様な生物種が存在
- さらにヒトでも個体ごとにも設計図の差がある



➡ 21世紀の最先端の研究対象

シミュレーションにも、データ解析にも
スーパーコンピュータの存在が欠かせない

DNA（デオキシリボ核酸）の構造



4つの部品（塩基）

C: Cytosine（シトシン）

A: Adenine（アデニン）

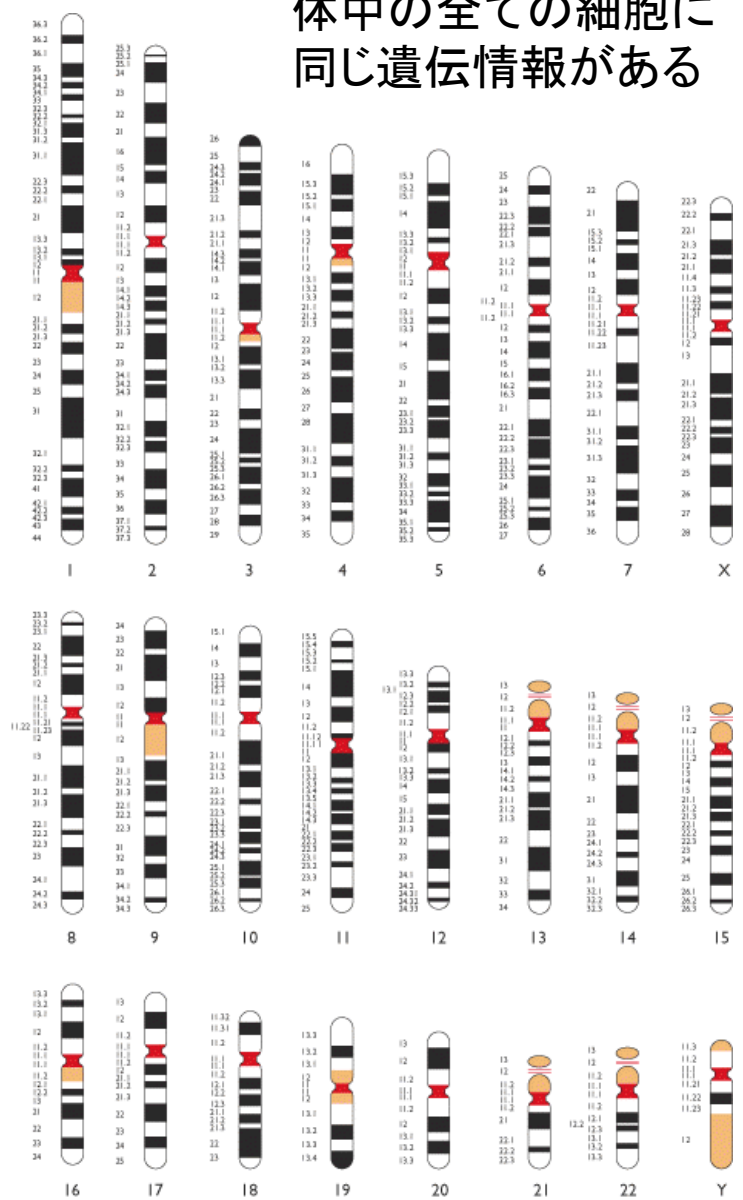
T: Thymine（チミン）

G: Guanine（グアニン）

4種類の文字でつづられた文章

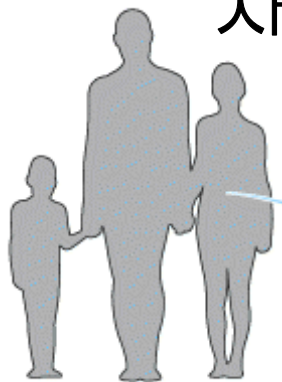
... **G** **T** **A** **C** ...
| | | |
... **C** **A** **T** **G** ...

体中の全ての細胞に
同じ遺伝情報がある

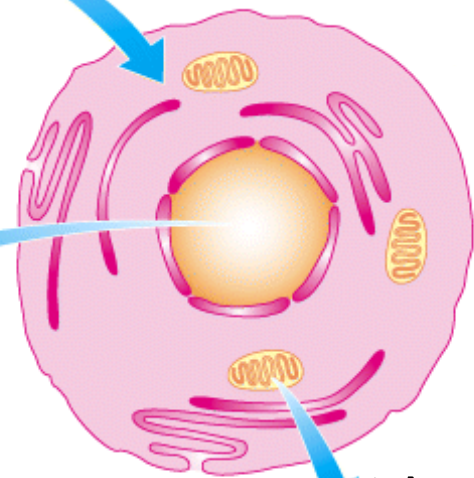


ヒトの染色体 1番～22番染色体
X染色体、Y染色体

人間・家族



ヒト細胞

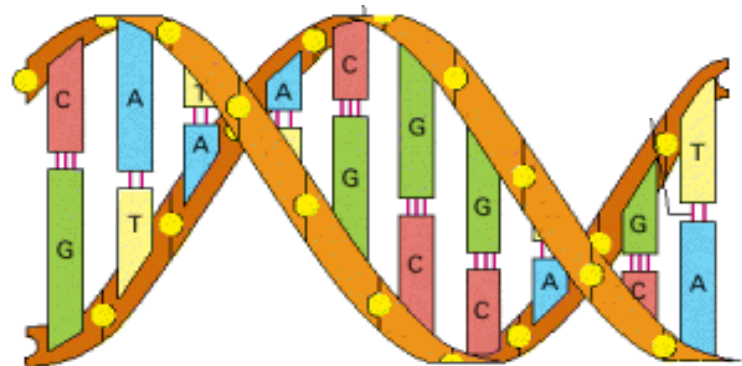


核ゲノム



=

ミトコンドリア
ゲノム



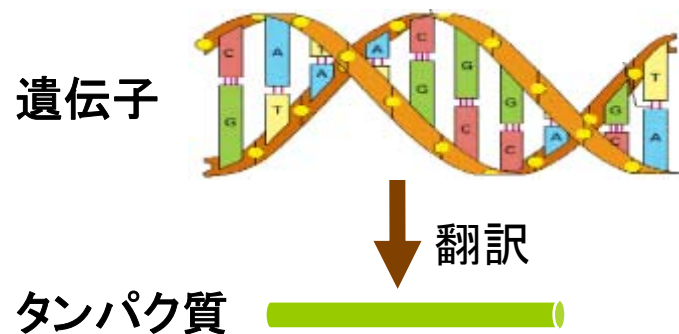
DNA
二重らせん

書籍“Genomes3”より引用

ゲノムと遺伝子は 同じ？ 違う？

ゲノム： 遺伝情報の総体 細胞内の全DNA
英語ではGenome (=Gene 遺伝子 + ome 総体)

遺伝子： 遺伝情報の要素単位 英語では Gene
あるタンパク質を合成するなど1つの役割に対応



必要な時に、必要な部分が、
必要なだけ読まれて、
細胞を支える部品や働き手
となるタンパク質を生産する。

**スーパーコンピュータ
でなければできない計算**



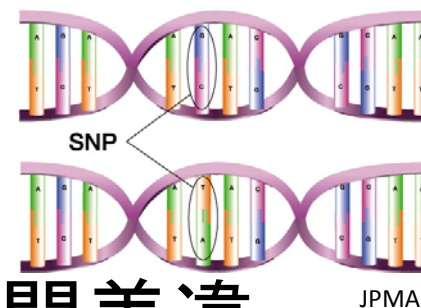
その1：膨大なパズルの解を探す

- 例・SNPsの組み合わせ**
- ・細胞内ネットワークの地図**

ゲノム配列の個人差

- **SNP** (Single Nucleotide Polymorphism)

「単塩基多型」と訳される。スニップ。
約1000塩基あたり1塩基程度の個人間差違。
ただし人口の1%以上の頻度で存在するもの。

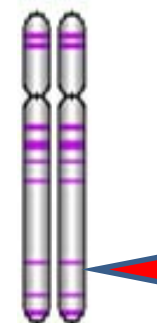


例： ヒト12番染色体 ALDH2遺伝子（お酒を飲んで赤くなるか？）

```
SQ Sequence 135 BP; 31 A; 25 C; 48 G; 31 T; 0 other;
caaattacag ggtcaactgc tatgatgtgt ttggagccca gtcacccttt ggtggctaca
agatgtcggg gagtggccgg gagttgggcg agtacgggct gcaggcatac actgaagtga
aaactgtgag tgtgg
```

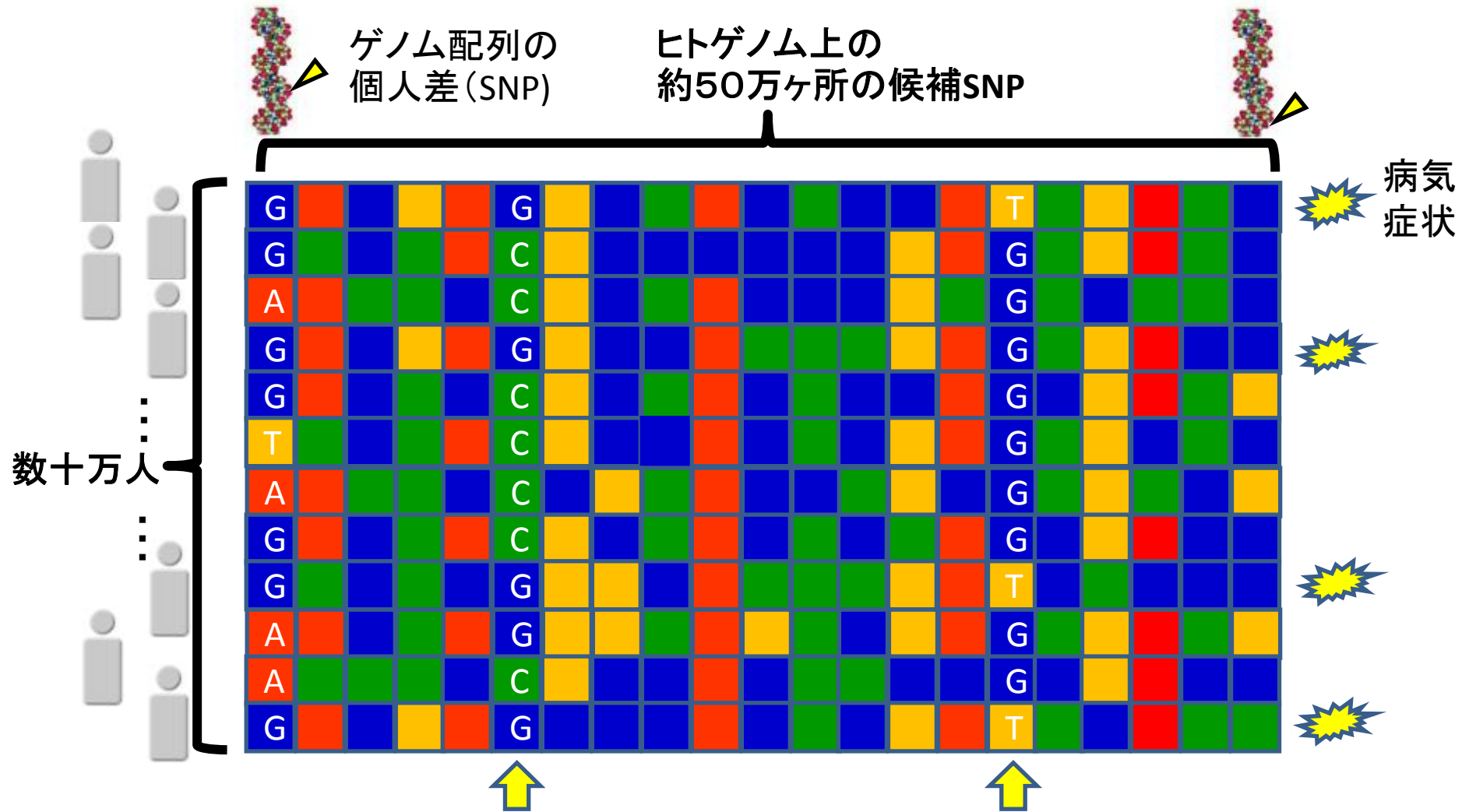
//

父母から受け継いだ2本の染色体がどちらも*2型だと酒に弱い。
アルコールが分解されてできるアルデヒドの処理に時間がかかる。
日本人では*1/*1型の人が約56%、*1/*2が約38%、*2/*2が約4%。



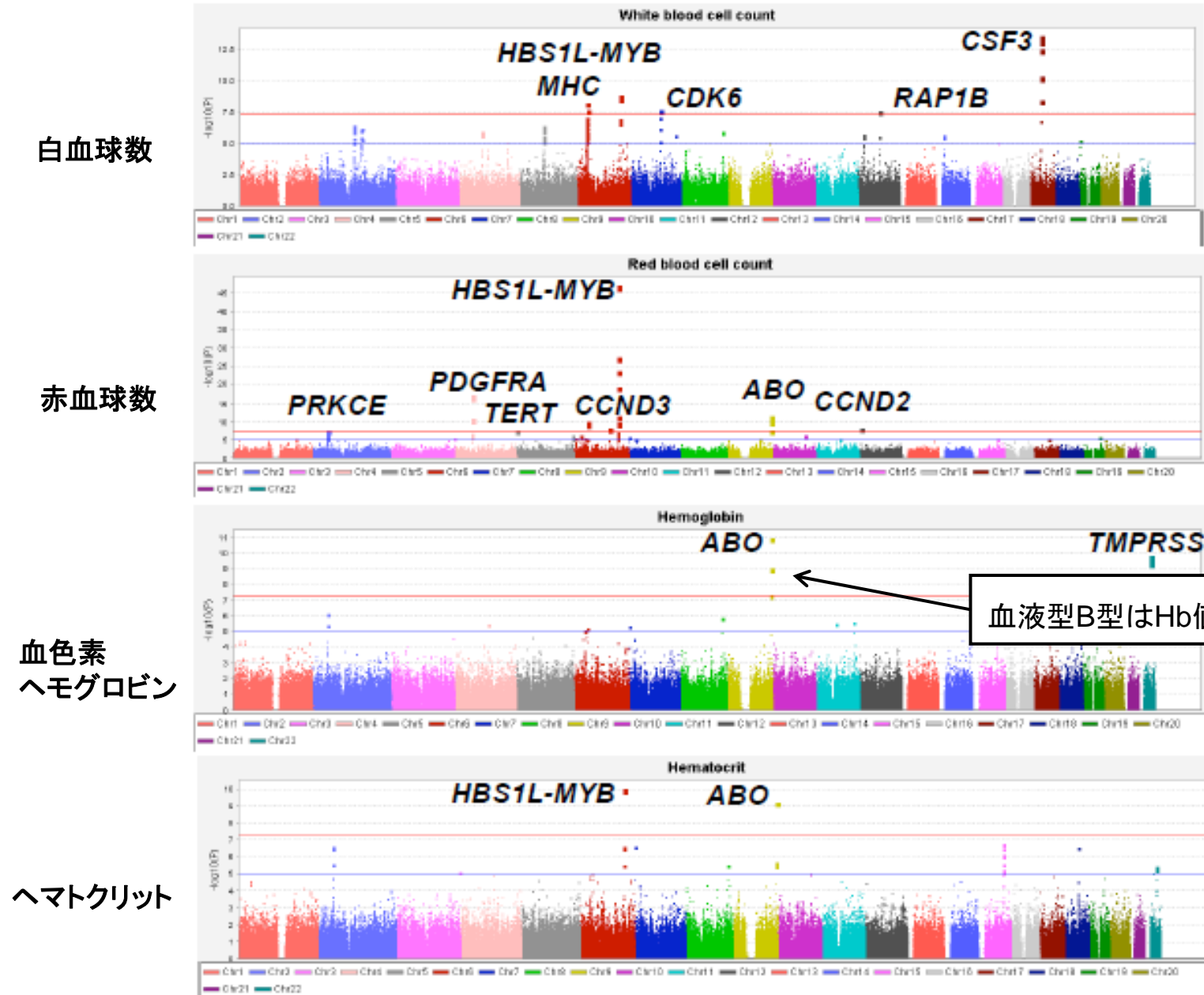
g (*1型)
a (*2型)

病気と連関するSNPを探したい



2箇所の組み合わせなら、 $(500,000 \times 499,999) / 2$ 通りの候補の作り方がある。
それぞれについて、(100%説明できることは無いので)確率的な評価計算を行う。

血液検査の項目の平均値がSNP型によって大きく変わる例を発見



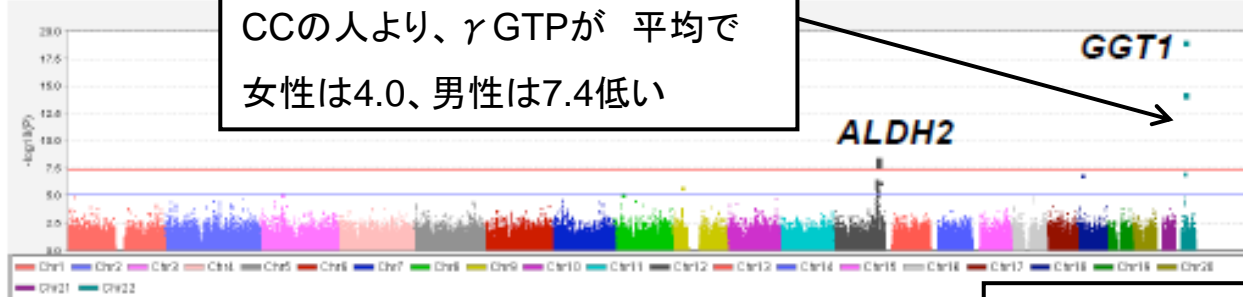
図の見方

偶然とは
考えられない
平均値のズレ

各染色体
上のSNP

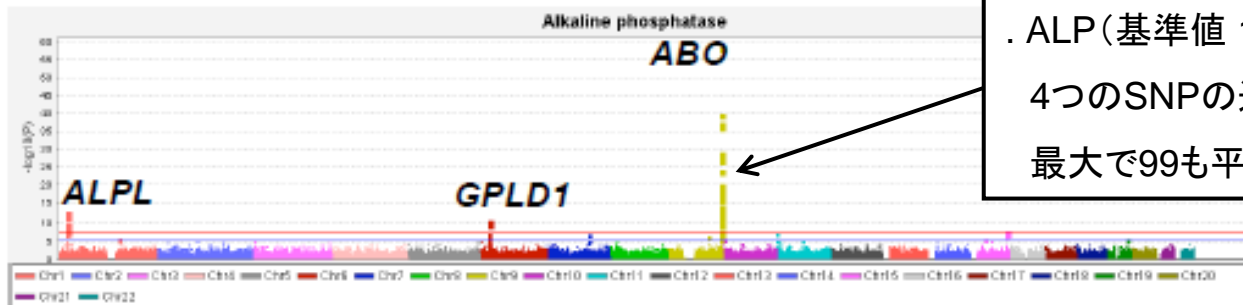
GGT
 γ-GPT
 肝機能

GGT1内のあるSNPがTTの人は
 CCの人より、γGTPが 平均で
 女性は4.0、男性は7.4低い

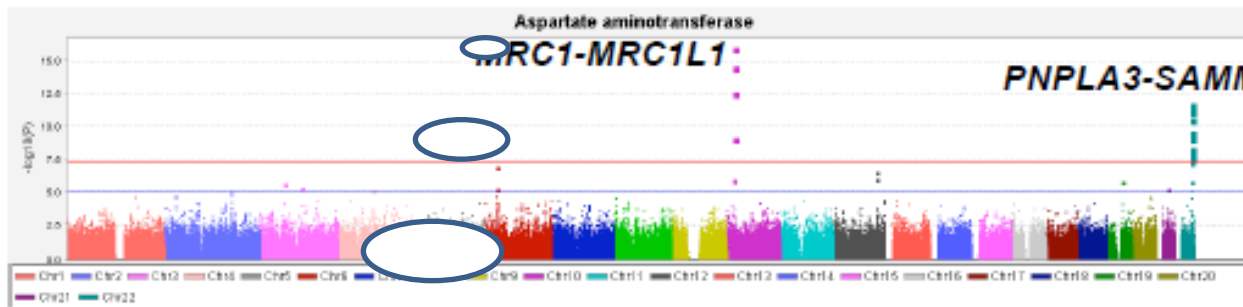


ALP
 胆道
 肝機能、骨

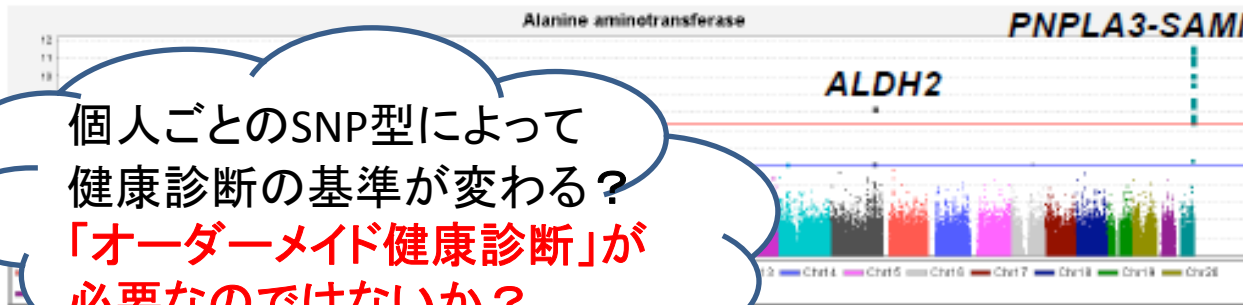
ALP(基準値 110-354 U/L)
 4つのSNPの遺伝型により
 最大で99も平均値が違う



AST

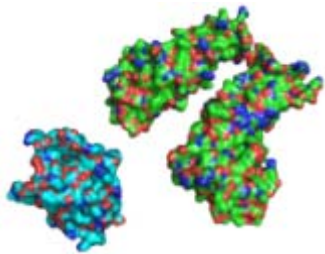
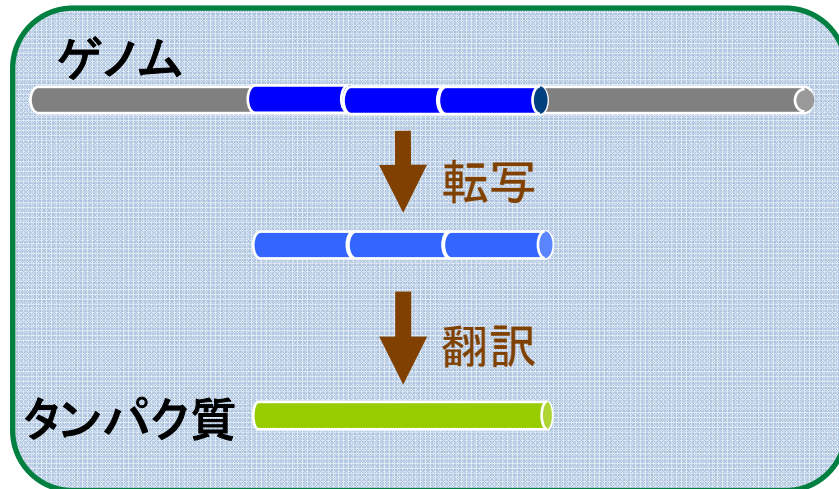


ALT



個人ごとのSNP型によって
 健康診断の基準が変わる？
 「オーダーメイド健康診断」が
 必要なのではないか？

細胞内のネットワーク図を知りたい

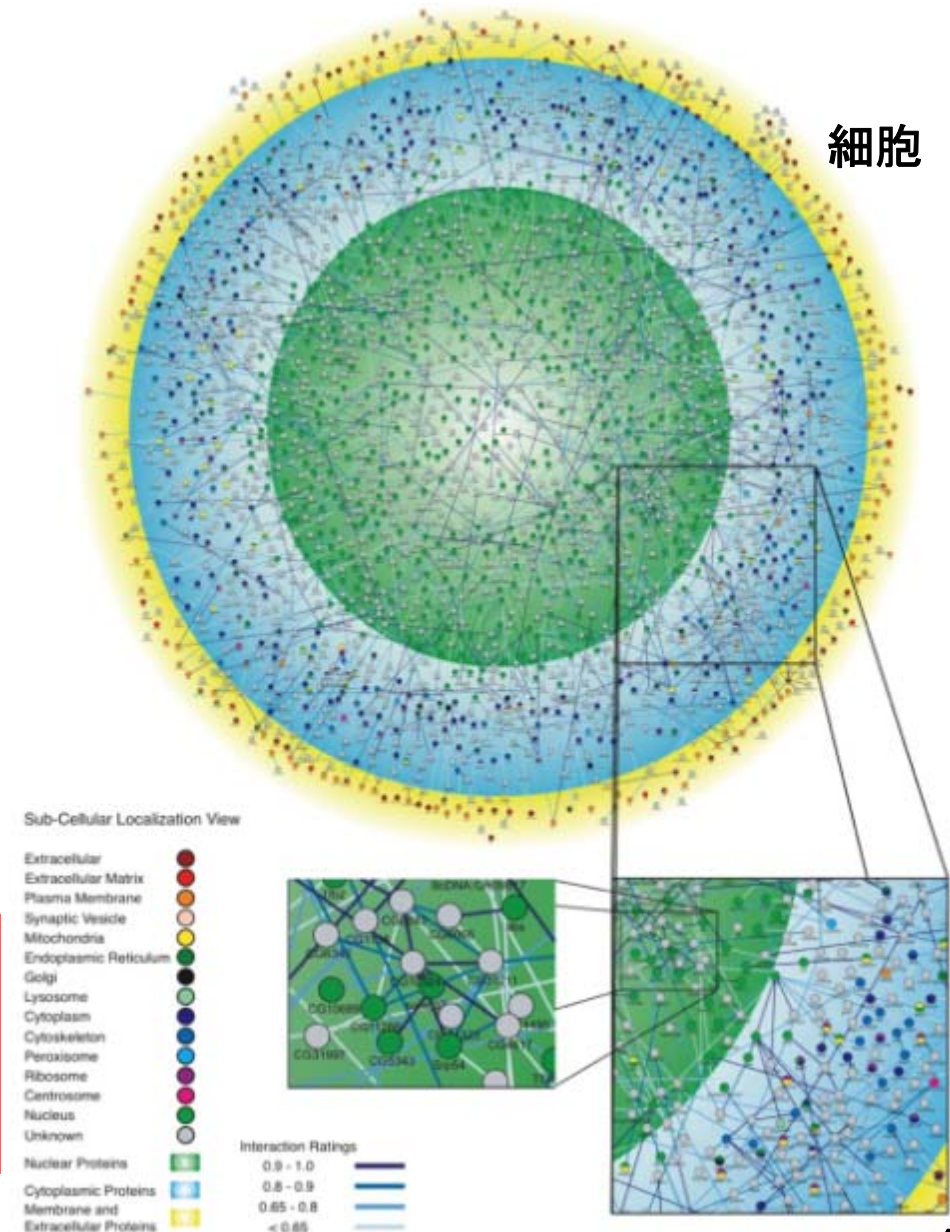


タンパク質は
ドッキングする。

相互作用の複雑な網。

限られた観測データから、ネットワークのつながり方や、その時点での信号の流れ方を「推定」するのは、さまざまな仮説を調べ尽くす必要がある複雑なパズル

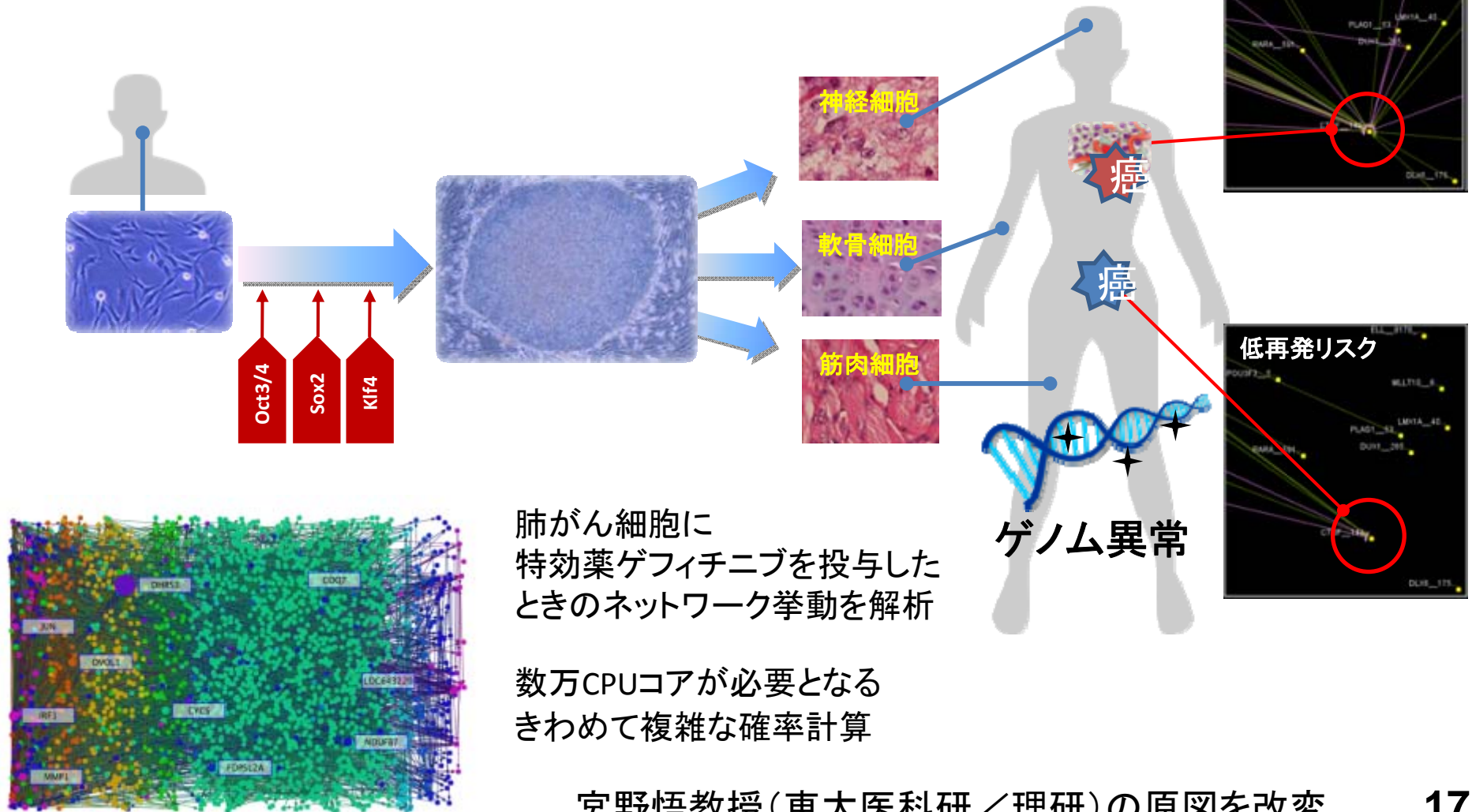
瀬々潤 准教授(東工大)の原図を改変



L. Giotら Science, 2003.

ネットワークの個人差から診断へ

がんの「個性」を、遺伝子ネットワーク全体として理解する。
がんの再発リスクの高さが評価できるようになってきた。。



**スーパーコンピュータ
でなければできない計算**



その2：精密なシミュレーション

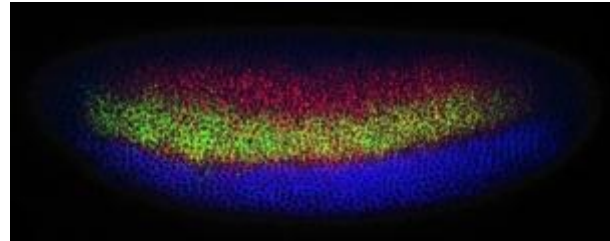
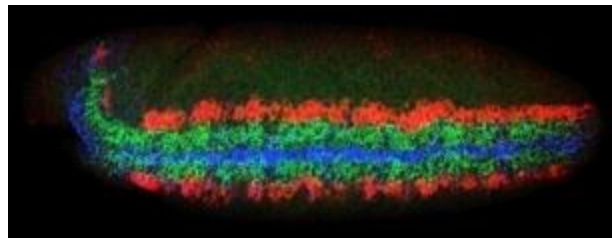
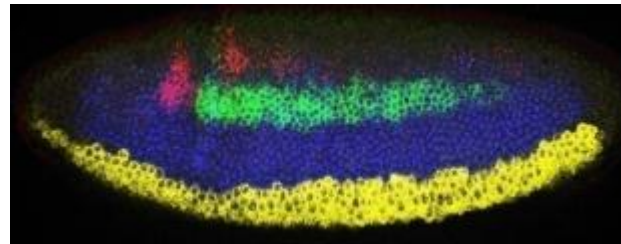
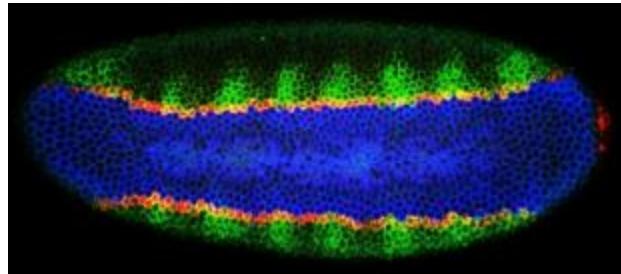
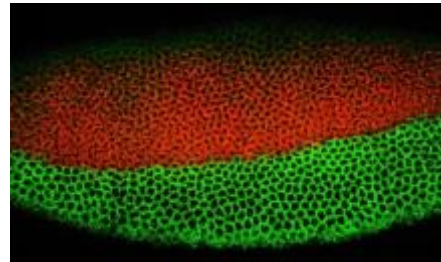
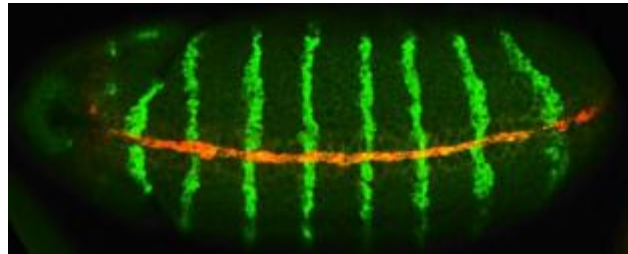
- 例・細胞の運命**
- ・薬剤の結合性評価**
- ・治療計画支援**

体が形作られていく不思議

全ての細胞は同じゲノムを持っているのに…

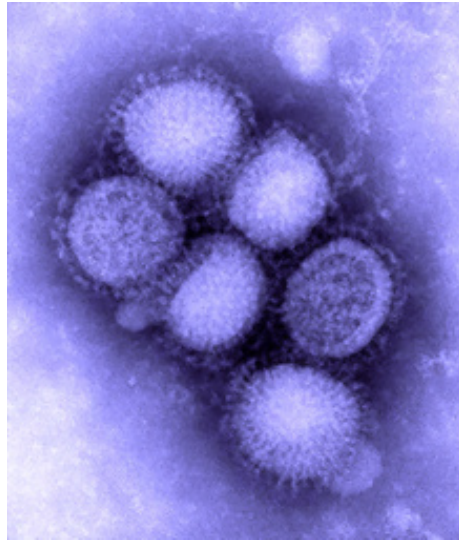
どうして細胞は役割を認識するのか(“細胞分化”という)。

物質の濃度等から細胞の運命を計算機で再現する試み。

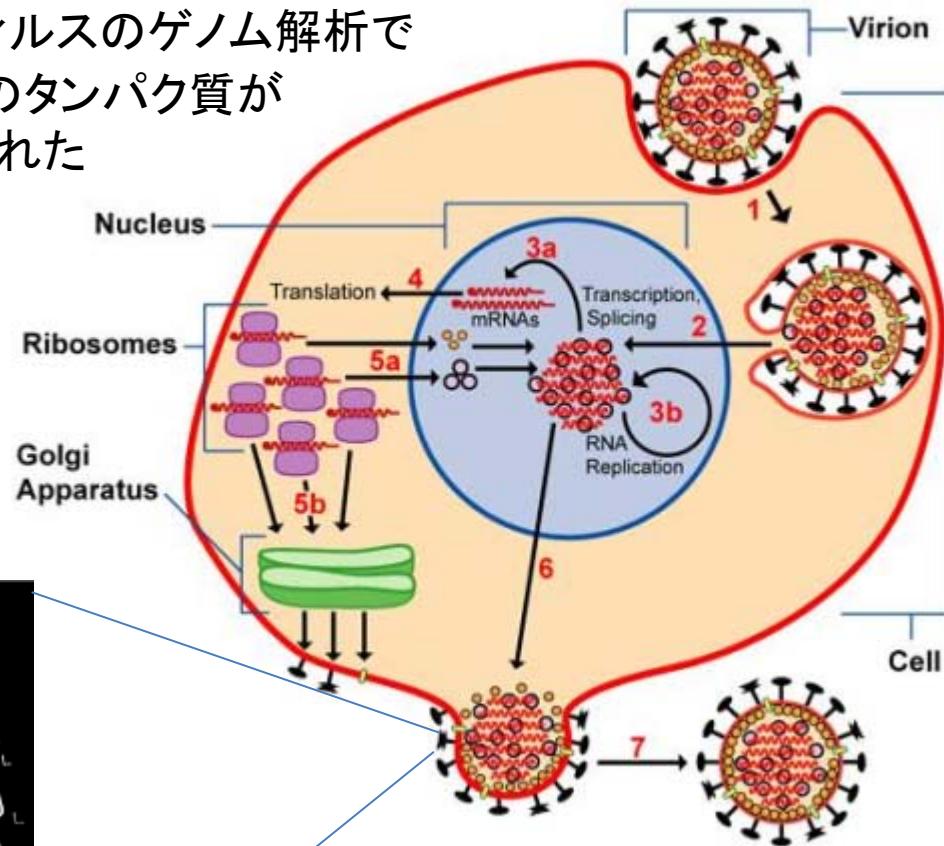


卵が細胞分裂
していく過程で、
場所ごとに
異なる遺伝子が
発現(コピー)
されている様子

抗インフルエンザウイルス薬の開発

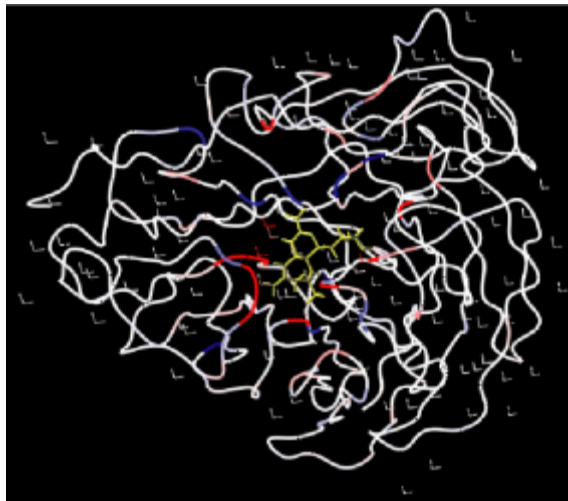


Flu ウィルスのゲノム解析で
10種のタンパク質が
発見された



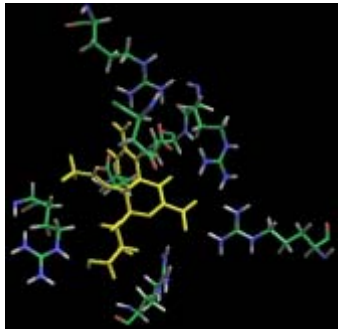
宿主細胞内
でのウイルス
増殖の機構
が解明された

NCBI Web
より引用



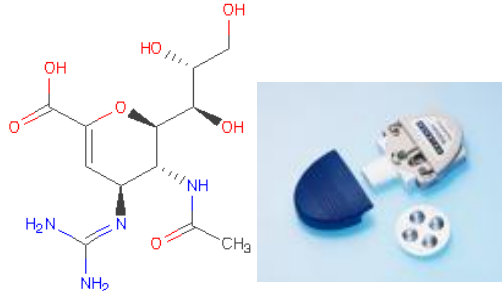
ノイラミニダーゼ酵素 (NA)

インフルエンザウイルスが作り出す
「ノイラミニダーゼ酵素」の機能を邪魔
できれば、たとえ細胞内で増殖しても
宿主の細胞膜から離れて移動できない



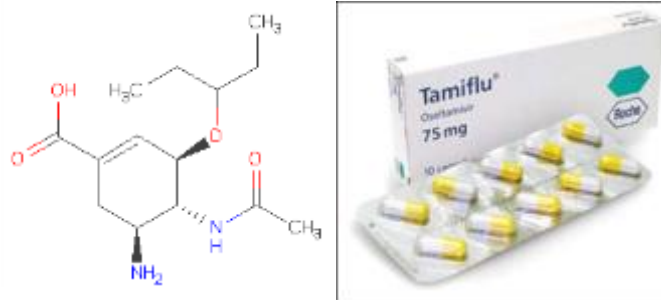
ノイラミニダーゼ酵素(NA) のくぼみ(ポケット)部分にぴったりはまり込み、NAの役割を邪魔(阻害)するような構造を持った化合物を、計算機上で、探す競争が世界中で繰り広げられた。

ザナミビル
(商品名 リレンザ)



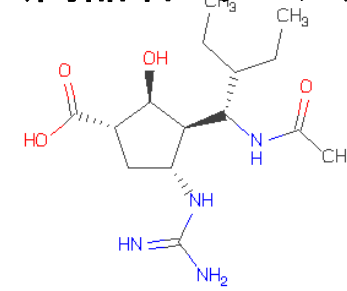
ビオタ社 (1989)

オセルタミビル
(商品名 タミフル)



ギリアド・サイエンス社 (1996)

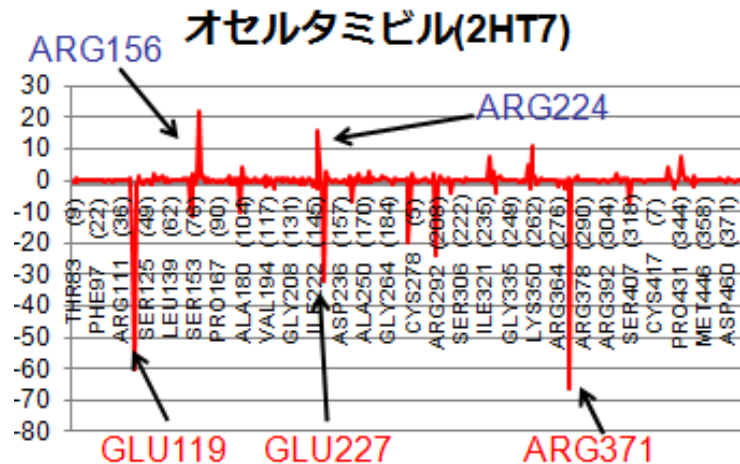
ペラミビル
(商品名 ラピアクタ)



バイオクリスト社

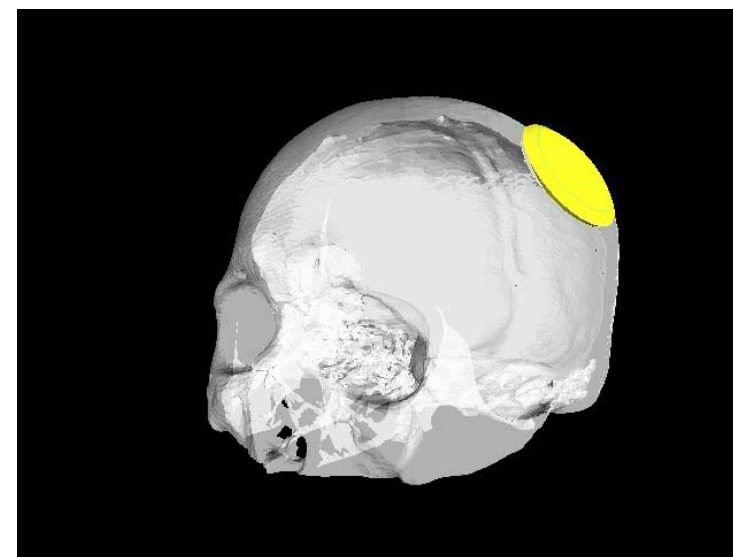
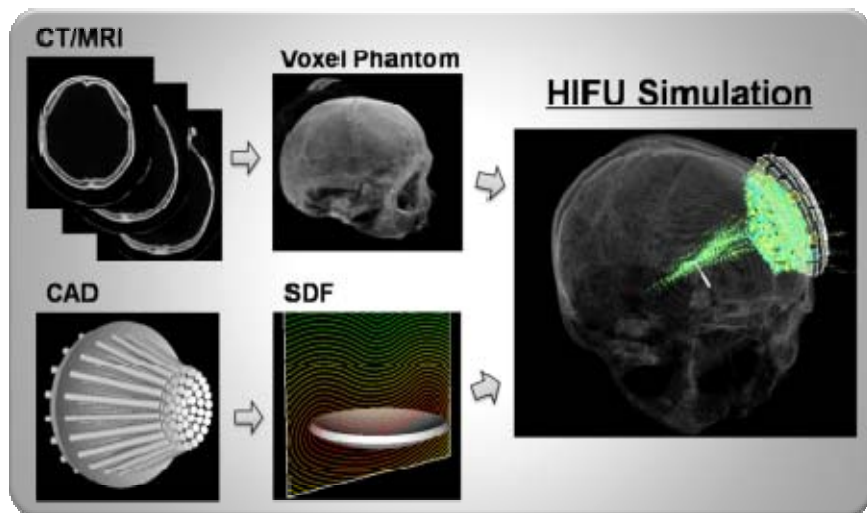
精密な「分子軌道法」計算の理論や、
並列計算のための技法は日本のお家芸！

しかし、あまりにも膨大な計算コストのため
従来までは創薬現場で使いにくかった。

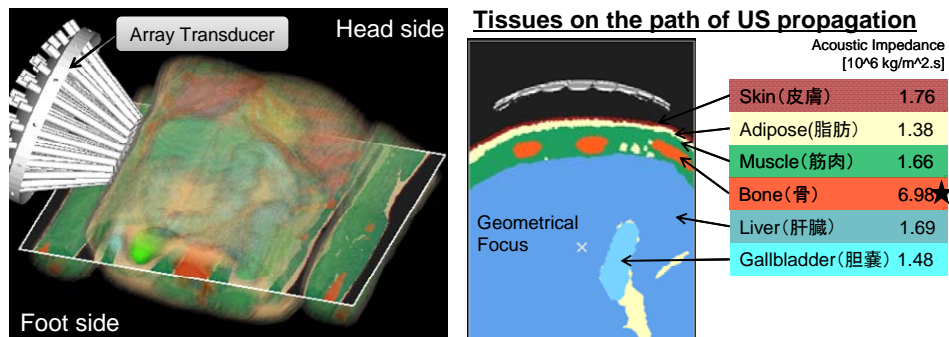


超音波治療器開発用シミュレーション

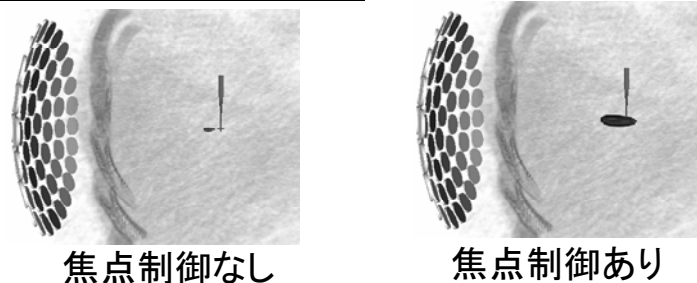
提供：
高木周博士(理研)



医用画像データとCADデータを直接利用するHIFUシミュレータ



国産初
実機設計のための
詳細シミュレーション



時間反転法による肝腫瘍焼灼シミュレーション

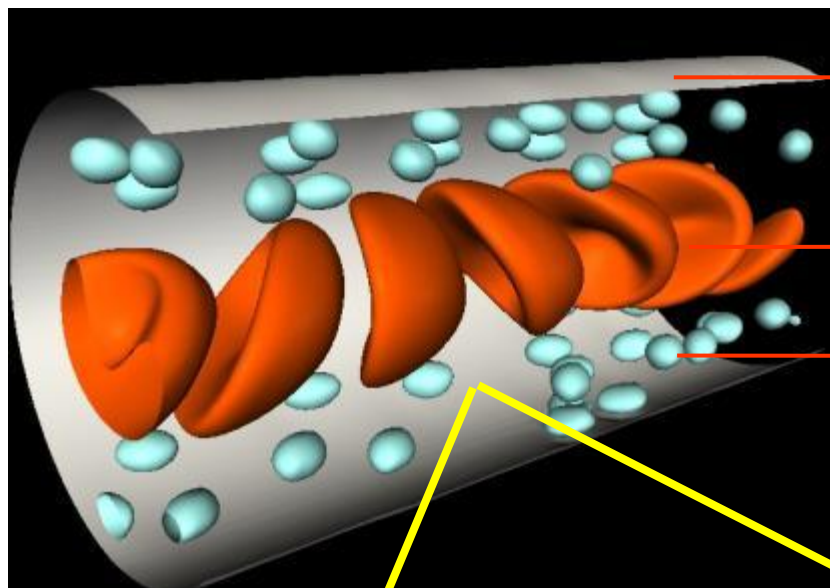
現状: 低解像度・ミリ秒スケール
超音波伝播シミュレーション

「京」では: 高解像度・実機設計
用詳細計算 & 秒スケール腫瘍
焼灼シミュレーション

血栓症の形成過程のシミュレーション

提供：
高木周博士(理研)

多数の赤血球や血小板を含んだ血流計算

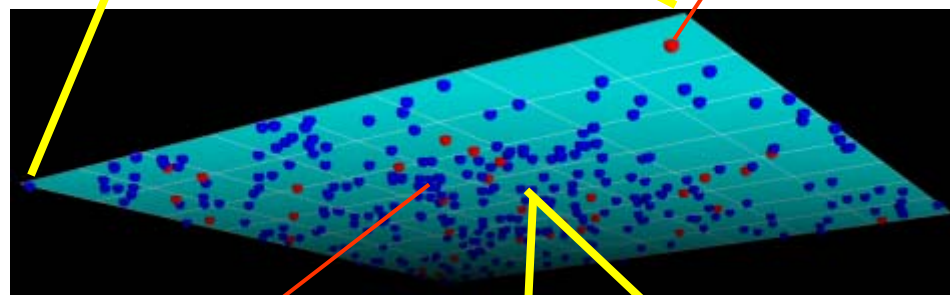


vessel wall

red blood cell

platelet

血小板, 血管壁間の分子間結合に関する動的モンテカルロ計算



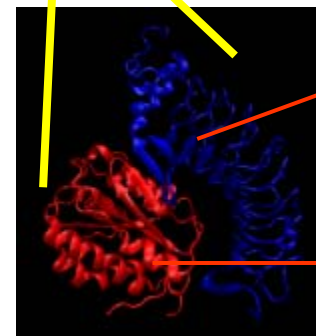
GPIb α -vWF bond

現状: 毛細血管が対象

「京」では: 心筋梗塞の対象となる直径数mm程度の冠動脈の計算が可能

GPIb α

リガンド-受容体分子の相互作用に関する分子動力学計算



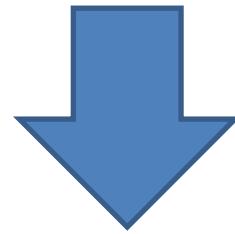
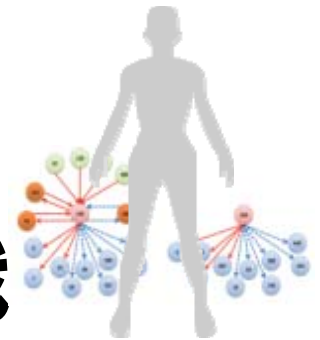
GPIb α N-terminal

vWF

A1 domain

まとめ

- 生命は「巨大システム」であり、その理解は21世紀の大きな挑戦



写真提供：(独) 理化学研究所

- スパコンでなければ不可能な計算
 - (1) 膨大なパズル (SNPs関連解析, ネットワーク推定)
 - (2) 精密なシミュレーション (細胞, 薬剤, 全身)