2013年12月20日　加古川東高校のみなさんへ
注：本スライドは加古川東高校での講義資料をもとに作成されています（2014年5月19日版）。

生物学とスーパーコンピュータ「京」
―講習―

理化学研究所

HPCI計算生命科学推進プログラム

江口至洋

yeguchi@riken.jp

今日は、こんなことをお話しします

- スーパーコンピュータ「京」って何？
  - 多くのCPUを協調的に働かせる仕組みを備えた、超並列電子計算機です。縁の下の力持ちとも言える人々の努力の賜物です。
- 生物学と化学、物理学、数学は陸続き！
  - スーパーコンピュータは多くの贈り物を私たちに届けてくれています。その一つに、「学問を結び付け、実りある協力関係を作り上げ、他の方法では解けない問題をも解く道を切り拓いてきた」ことがあります。生物学を例に、見てみます。
- 実習（なぜ生物学にスパコンが必要？）

スーパーコンピュータ「京」って何？

昔、計算機は特殊なもの。

計算機ENIAC（1945年）

アメリカ陸軍が大砲の弾道計算を目的にペンシルバニア大学と共同開発

Courtesy of Oak Ridge National Laboratory, U.S. Dept. of Energy

プログラム内蔵型計算機
EDVAC（1951年）

アメリカ陸軍とペンシルバニア大学との共同開発

今や、空気のようなもの。

身の回りは計算機にあふれています。
みなさんは計算機なしには生きていけない！

スーパーコンピュータ「京」？

どちらかというと、電卓やパソコンが巨大になったのではなく、多くのパソコンが相互に繋がって協力して働いています。

「京」の構成

システム全体

10PFLOPS

約8万個のCPUを相互に接続
50m×60mの部屋に鎮座

計算ラック
システムボード×24

12.3TFLOPS

システムボード

CPU

演算性能128GFLOPS

「京」の超並列計算を支える人々

20万本のケーブル、総延長1,000Km

世界のスパコン計算性能ランキング

順位	システム名称	設置場所／ベンダー	国名	速さ（PFLOPS）
1	天河2号	国防科学技術大学/Intel	中国	33.9
2	Titan	オークリッジ国立研究所／Cray	米国	17.6
3	Sequoia	ローレンス・リバモア国立研究所／IBM	米国	17.2
4	京	理研AICS／富士通	日本	10.5
5	Mira	アルゴンヌ国立研究所／IBM	米国	8.6

2013年11月

Flops（フロップス）一秒間に処理できる浮動小数点計算の数

スパコンの性能向上とみなさんのパソコン

みなさんが普段使っているパソコンは、20年ほど前の世界最速スーパーコンピュータと同等以上の性能を持っています。スーパーコンピュータも意外と身近にあります。

http://www.top500.org/statistics/perfdevel/

スーパーコンピュータの贈り物！

- **実験を助け、補完します**
  - ヒトゲノムのビッグデータを解析します
  - コンピュータなしに新しい自動車を設計することはありません
- **実験できないものを、「実験」します**
  - 計算機顕微鏡でタンパク質の動きを見ます
  - コンピュータが津波の襲来を予測します
- **人間の予測能力を向上させます**
  - 今やスーパーコンピュータなしに天気予報はできません
  - 心臓の動きをシミュレーションします
- **人間の思考能力を超える？**
  - 計算機と人間の知的な対話が数学の難問を解きました
  - コンピュータチェスが世界チャンピョンに勝ちました
- **新しい発見をもたらします**
  - ソリトン（孤立波）を見出す
- **学問を結びつけ、学際性を強めます**

空気抵抗の低減

気象庁による24時間予想図
（平成 25年10月5日9時の予想）

四色問題

葛飾北斎
『神奈川沖浪裏』1831年頃
http://ja.wikipedia.org/

スーパーコンピュータは、学問を結びつけ、学際性を強めます

例：　生物学と化学、物理学、数学は陸続き！

www.tmd.ac.jp/artsci/biol/textbook/profile.htm

離れ小島を繋げ出した人たち

生物学は離れ小島？

1965年　ワトソン 「遺伝子の分子生物学」
1958年　ケンドルー：　タンパク質のX線結晶構造解析
1953年　ワトソン、クリック：　DNAの二重らせんモデル

1865年　メンデル：　遺伝の法則
1859年　ダーウィン 「種の起源」

1735年　リンネ 「自然の体系」

1665年　フック 「ミクログラフィア（微小世界の図説）」

いや。意外と繋がっているのでは

生物は進化し、今のヒトがいる！

（1809~1882）

「自然選択、すなわち
生存闘争において有利な変種が保存されること
による種の起源について」

ところで、遺伝的変異がどのように生じ、いかにして子孫に伝えられるかという**遺伝の法則はどうなっているのだろう？**

http://darwin-online.org.uk/

おそらく粒子の形をした「遺伝子」があるのだろう、遺伝の仕組みを明らかにしよう！

メンデル

雑種植物の研究

岩槻邦男・須原準平訳

栽培植物の新品種を作るための人工交配をヒントに，メンデル（1822－84）は，ブルノの修道院の庭で8年にわたりエンドウの交配実験を行なった。きわめて科学的に遂行された実験を厳密に検証したこの論文は，当時はその価値を認められず，1900年になって再発見され，遺伝学の基本的な法則にメンデルの名が冠せられることとなった。

青932-1
岩波文庫

（1822~1884）

２年かけて、純系種を選出し、８年かけて交配実験を行うという「**かなりの勇気を必要とする実験**」を行った。

優性形質Aと劣性形質aをもつ雑種第１世代（F_1）**Aa**を掛け合わせると、雑種第２世代（F_2）に対して次の展開式が得られる。

(A+a)×(A+a)＝AA+2Aa+aa

私は実験物理学の教師です

http://www.jic.ac.uk/germplas/pisum/zgs4f.htm

ワトソンとクリック以前（1953年以前）

ヒト細胞の
核内にある染色体

G-band 染色体

女性と男性では異なります。この図はどちら？

1953年以前までの一般的な考え
「遺伝子は核内の染色体の中のタンパク質が担っているのだろう。」

ワトソンとクリックが言ったこと！

1953年　ワトソンとクリック
「遺伝情報はDNAの塩基配列に刻まれている。」

PDB：3BSE

1965年
ワトソンの著書「遺伝子の分子生物学」
第2章　細胞は化学の法則に従って生きている

さまざまな生命現象に対するわれわれの理解を**分子レベル**にまで掘り下げていくならば、やがてわれわれは、生きていく状態の基本的な性質を理解しうる。

20世紀後半の生物研究の大きな流れを作った主張です
「さまざまな生命現象に対する私たちの理解を分子レベルにまで掘り下げていく」

Watson JW, Crick FHC (1953) Nature, **171**, 737

タンパク質の機能と構造！

1836年にシュワンによって発見された酵素ペプシン

機能：
胃の主細胞から分泌される。
胃の中の酸性環境下で食物を消化するよう最適化されている。
非常にでたらめな酵素で、タンパク質を様々な長さの断片に切断する。

この機能を構造から理解しよう

http://www.pdbj.org/mom/index.php?l=ja&p=012

花成ホルモン（フロリゲン）！ 1936年チャイラキアンが提唱

2008年版の高等学校の教科書

「花成ホルモン：　光周性を示す植物の花芽形成の促進にはたらく物質は花成ホルモン（フロリゲン）と名付けられたが、その実体は明らかではない。」

⇒この記述で重要なこと

花成ホルモンの機能（花芽形成の促進）をその実体、つまり分子のレベルで明らかにしようとしています。

2008年版
「リードα　生物Ⅰ」
数研出版編集部　編

花成ホルモン（フロリゲン）！

花成ホルモン（フロリゲン）の分子構造を明らかにする。

PDB（3AXY）

計算機と生物学の繋がりが強くなってきました

2012年　スーパーコンピュータ「京」

2005年　次世代DNAシークエンサー

1977年　サンガー：　DNA塩基配列決定法

1965年　ワトソン「遺伝子の分子生物学」
1958年　ケンドルー：　タンパク質のX線結晶構造解析
1953年　ワトソン、クリック：　DNAの二重らせんモデル

1865年　メンデル：　遺伝の法則
1859年　ダーウィン「種の起源」

1735年　リンネ「自然の体系」

1665年　フック「ミクログラフィア（微小世界の図説）」

計測技術と計算機の飛躍的進歩は、生物学と物理、化学が陸続きであることをわたしたちに気付かせてくれています。

「コンピュータシミュレーション」は試験問題にも出ています

2006年の大学入試センター試験「理科 ③」において、以下のような問題が出題されている。

第2問　次の文章を読み，下の問いに答えよ。

A　**コンピュータの利用について考えよう。**
半導体技術の進歩により計算処理能力も高まったので，コンピュータで仮想的に実験を行い，大規模で複雑な現象を予測することもできるようになった。

問2　上の文章中の下線部を表す語として最も適当なものを，次の①～⑤のうちから一つ選べ。

① コンピュータシミュレーション
② コンピュータグラフィックス
③ CT(コンピュータトモグラフィー)
④ コンピュータ制御
⑤ コンピュータ計測

ヒトゲノムの塩基配列（30億文字）を全て読み取る

ヒトゲノム

イルミナ（株）のご厚意による

ヒトの場合30億塩基対

１番染色体	ATCCATCT・・・TCCCTTCTCTG
２番染色体	GAGACTTG・・・GGAGCC
	・・・・・・・
	・・・・・・
Y染色体	GAACCATG・AAG

gtgccacctc ctgggacgct gccctcgggg tccccgatcg cgaccaaaca agtaagagaa ctctgttgaa gctcgatcct cacctactct ctaaataaaa gaatattcca gagaaggaaa gagagaagca caaatattaa gacttctaga
cccctggaga ggactaaagt ctacccctcc acatggcttc ttctccaatg tgatgctgga gttgaatgaa cttaggactc cactccaagg ttgacttgaa aggtgtgtag aagattggaa agagaaaggt attaagtcat gatgcaggaa
tctgcgacag agacaataag caacagttca atgaatcaaa atggaatgag cactttaagc agccaattag atgctggcag cagggatggg agatcaagtg gtgacaccag ctctgaagta agcacagtgg agctgctgca tctgcaacaa
cagcaggctc tccaggcagc aagacaactt ttgttacagc agcaaaccag tggattgaaa tctcctaaga gcagtgagaa gcagaggcca ctgcaggtgc ccgtgtcggt ggccatgatg actccccaag tgatcacccc tcagcagatg
cagcagattc ttcagcagca ggtcctgtct cctcagca[illegible] tccaggccct cc[illegible]cagcag cagcaag[illegible] tg[illegible]gttgca gcagcaacaa ctacaag[illegible] tttacaag[illegible] aca[illegible]aagaa cagttacat[illegible] ttcagctttt [illegible]gcaac[illegible]cag
caacag[illegible]gcagcaaca[illegible]gcag[illegible]ag cagcagc[illegible] ag[illegible]gca[illegible]a acagcag[illegible]ag cag[illegible]gcaac[illegible]acagc[illegible]gcag[illegible]aa cagcagc[illegible] a[illegible]gcaaca[illegible]ca[illegible] aa[illegible]aagcaa [illegible]agcagc[illegible]gcagc[illegible]cag
cagcagcagc agttggc[illegible]gc g[illegible]gcagc[illegible]c gtctt[illegible]cag[illegible]agcagcttct ccaa[illegible]gcag cag[illegible]ac[illegible]ag[illegible]acaa[illegible] t[illegible]gctca[illegible] cttca[illegible]cgc[illegible]agggcct[illegible]t ctc[illegible]tccca c[illegible]ggccaag [illegible]agcccttcc tgtc[illegible]gtcg
ctgcctcaag ctggctt[illegible]ag [illegible]cctgctg[illegible] attcag[illegible] tatggaaaga ag[illegible]tactg[illegible] gt[illegible]ca[illegible]ta[illegible]g[illegible]aga[illegible]aa [illegible]ggcatca[illegible] catgga[illegible] [illegible]tagacctc[illegible] gac[illegible]acaat tc[illegible]ctcga [illegible]tacctcctc ca[illegible]cacgtcc
aaa[illegible]catcac cacccat[illegible]acatcat[illegible]cc atagtga[illegible] gacagtcttc ag[illegible]tctgaat gcaag[illegible]ggg[illegible]c[illegible]ctc[illegible]c[illegible]acatga[illegible]g actggg[illegible] cgcaca[illegible]ct ct[illegible]tgg[illegible]at g[illegible]tct[illegible] [illegible]tggccagg ct[illegible]tgaaagc
atatg[illegible]cttggaca gtt[illegible]tgaag cacc[illegible]aaca [illegible]gcatgc [illegible]ggatgac cg[illegible]gcactg [illegible]atgtcg agt[illegible]aaatg cagg[illegible]ggtac [illegible]gtt[illegible] [illegible]acagc [illegible]aagg[illegible] gcgaacgtct t[illegible]gcgatg
atgacccact tgcacatgcg accctcagag cccaaaccat ctcccaaacc tctaaatctg gtgtctagtg tcaccatgtc gaagaacatg ctggagacat ccccacagag cttacctcaa acccctacca caccaacggc cccagtcacc
ccga[illegible]taccc [illegible]ggggccctc tgtaat[illegible]acc ccagccagtg tg[illegible]ccatgt gggagccata cgaaggcgac attcagacaa ata[illegible]aacatc cc[illegible]atgtcat cagagattgc cc[illegible]aactac gaattttata aaaacgcaga tgttagacct
cc[illegible]tt[illegible]gcacact cata[illegible]gc[illegible]gtctcat[illegible]gcca[illegible]aggcagtta [illegible]cttaag a[illegible]atttag ct[illegible]tcca[illegible]ta[illegible]g[illegible]atgcc[illegible]gcactg[illegible]agaa[illegible]gt gcgtcataat
cttag[illegible]tgc[illegible]aaatgttt tgttc[illegible]gaa[illegible]atgtt[illegible]aggag[illegible]tggactg[illegible]gaagaagt[illegible] aatac[illegible]agaa[illegible]gcaagg[illegible]ca[illegible]agata[illegible] cagg[illegible]c a[illegible]ttgta[illegible]aaatatt[illegible] ctac[illegible]ttt [illegible]gctttga
gcagctctaa atgccagtct gcaggctgcc ttggcagaaa gcagcttgcc tttgctaagt aaccctggac tgatcaataa cgcctccagt ggcctcctgc aggccgtcca cgaagacctc aatggttccc tggaccacat cgacagcaat
gggaacagca gtccgggttg ctcacctcag ccacacatac attcaatcca tgtcaaggaa gaacccgtga ttgcagagga tgaagactgt ccaatgtcat tagtgacaac agccaatcac agtccggagt tagaagatga cagagagatt
gaggaagagc ctttatctga ggacctggaa tgagaacgaa cttgtgacac ctcagtgtga agggacatat cactgacctt cataaccact ccacaaccat gaatagttga caaattttta ctgtgactat ttattaagca tggataaagg

この文字列の文法は？

今だ、全てが解っているわけではありません。

www.riken.jp/r-world/info/release/press/2004/040803/index.html

ヒトゲノムの塩基配列（30億文字）を読み取るお値段

Adapted from the National Human Genome Research Institute, USA

Copyright © 2014 Illumina, Inc.. Used under permission.
Copyright © 2014 Life Technologies Corporation. Used under permission.

赤血球と血小板を含む血流のシミュレーション

人の血液中に占める赤血球の体積の割合はほぼ35～50%です（女性は少な目です）。
赤血球は柔らかく、血管中でくにゃくにゃと変形しながら流れています。
白い物体は血小板です。赤血球に押され、壁際を流れます。

心臓の動きをシミュレーション（計算機ビデオ）

心電図

・心筋細胞内のたんぱく質の確率的運動から細胞の収縮、心拍動、血液駆出、冠循環までを一貫してシミュレート。
・ シミュレーションから超音波エコー、流速ドップラー、心電図、カテーテル検査などの精緻なデータが再現される。そのデータを基に病態の解析が可能に。

26

「科学の精密化」と「科学の統合化」

Life Technologies

Illumina

DNAシークエンサー

NMR（理研）

Spring-8（理研）

計測技術の飛躍的進歩

生物学、医学のビッグデータ

ビッグデータの多階層統合

生命を理解し、予測するための

細胞

組織

心臓

Homo sapiens
（約3万遺伝子）

もはや、生物学と物理学、化学、数学は、離れ小島ではなく、陸続きになっています。

Copyright © 2014 Illumina, Inc.. Used under permission.
Copyright © 2014 Life Technologies Corporation. Used under permission.

2013年12月20日　加古川東高校のみなさんへ
注：本スライドは加古川東高校での講義資料をもとに作成されています（2014年5月16日版）。

生物学とスーパーコンピュータ「京」
ー実習ー

理化学研究所

HPCI計算生命科学推進プログラム

江口至洋

yeguchi@riken.jp