

クレジット:

UTokyo Online Education 学術俯瞰講義 2019 井元清哉

ライセンス:

利用者は、本講義資料を、教育的な目的に限ってページ単位で利用することができます。特に記載のない限り、本講義資料はページ単位でクリエイティブ・コモンズ 表示-非営利-改変禁止 ライセンスの下に提供されています。

<http://creativecommons.org/licenses/by-nc-nd/4.0/>

本講義資料内には、東京大学が第三者より許諾を得て利用している画像等や、各種ライセンスによって提供されている画像等が含まれています。個々の画像等を本講義資料から切り離して利用することはできません。個々の画像等の利用については、それぞれの権利者の定めるところに従ってください。

東京大学 教養学部 前期課程 2019年度S Semester
主題科目 学術フロンティア講義
学術俯瞰講義

「新しい医療が社会に届くまで～データサイエンスが支える健康社会～」

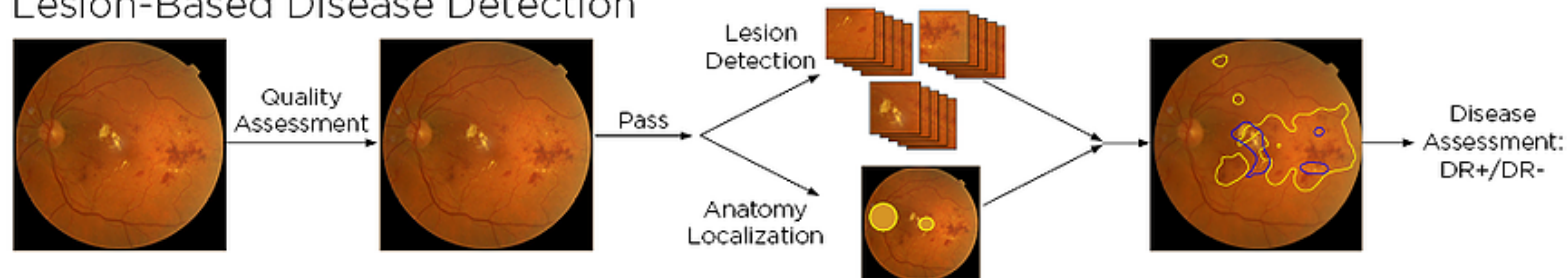
ゲノム情報と人工知能を用いた 新しい時代のがん医療

井元 清哉

東京大学医科学研究所
ヘルスインテリジェンスセンター
健康医療データサイエンス分野

糖尿病性網膜症機器IDx-DrをFDAが承認(2018 4/12)

Lesion-Based Disease Detection



左右2枚の眼底写真を撮影し、クラウドサーバーに送信すると1分以内で**初期**糖尿病性網膜症が診断される。

感度87%特異度90%。

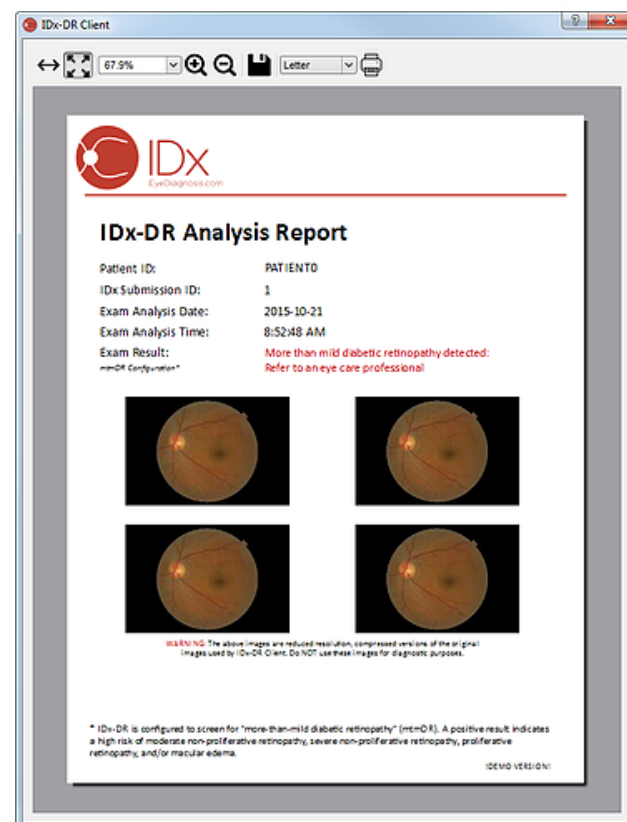
網膜症あれば眼科専門医受診、なければ1年後再受診。

開業医向け。

IDx Lesion-Based Disease Detection

IDx IDx-DR Analysis Report

<https://www.eyediagnosis.co/>



手首骨折診断OsteoDetectをFDAが承認(2018 5/24)

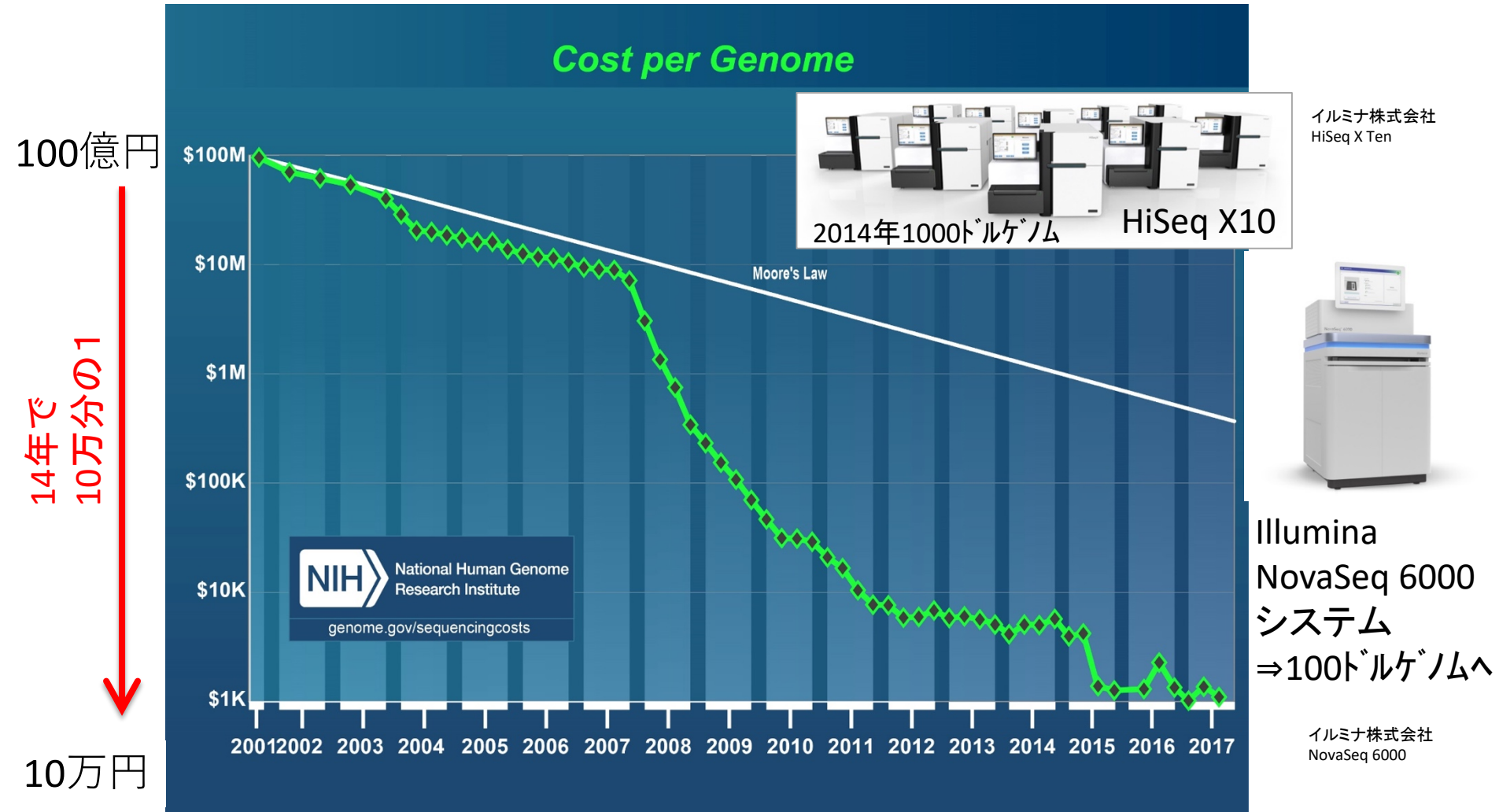


Photo by Tom Page, from flickr
<https://flic.kr/p/GUzz>
CC BY-SA 2.0

手首(橈骨遠位端)骨折を検出するアルゴリズム搭載ソフトウェア。2枚X線写真の骨折想定箇所に印を表示。約1,000件のX線画像を教師データとする機械学習。ソフト併用で、整形外科医単独より感度特異度向上。診断支援。開業医+α向け。

<https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm608833.htm>

一人分のゲノムシーケンスを得るコスト (2018年1月の入札では5～6万円)



Courtesy: National Human Genome Research Institute



ゲノムビッグデータの時代



集団学のパワー。最強のスループット。1000ドルゲノム。

HiSeq X Tenシステムは、10台の超ハイスループットシーケンサーから構成される設備で、大規模なヒト全ゲノムシーケンスを目的としてデザインされました。



Illumina HiSeq X-ten
1万8千人の全ゲノム解析/年
1000ドル/人
米国、英国、オーストラリア
中国、韓国、カタール

イルミナ株式会社
HiSeq X Ten



President Barack Obama
Precision Medicine Initiative
2015/1/30
約 250 億円の資金、100 万人以上

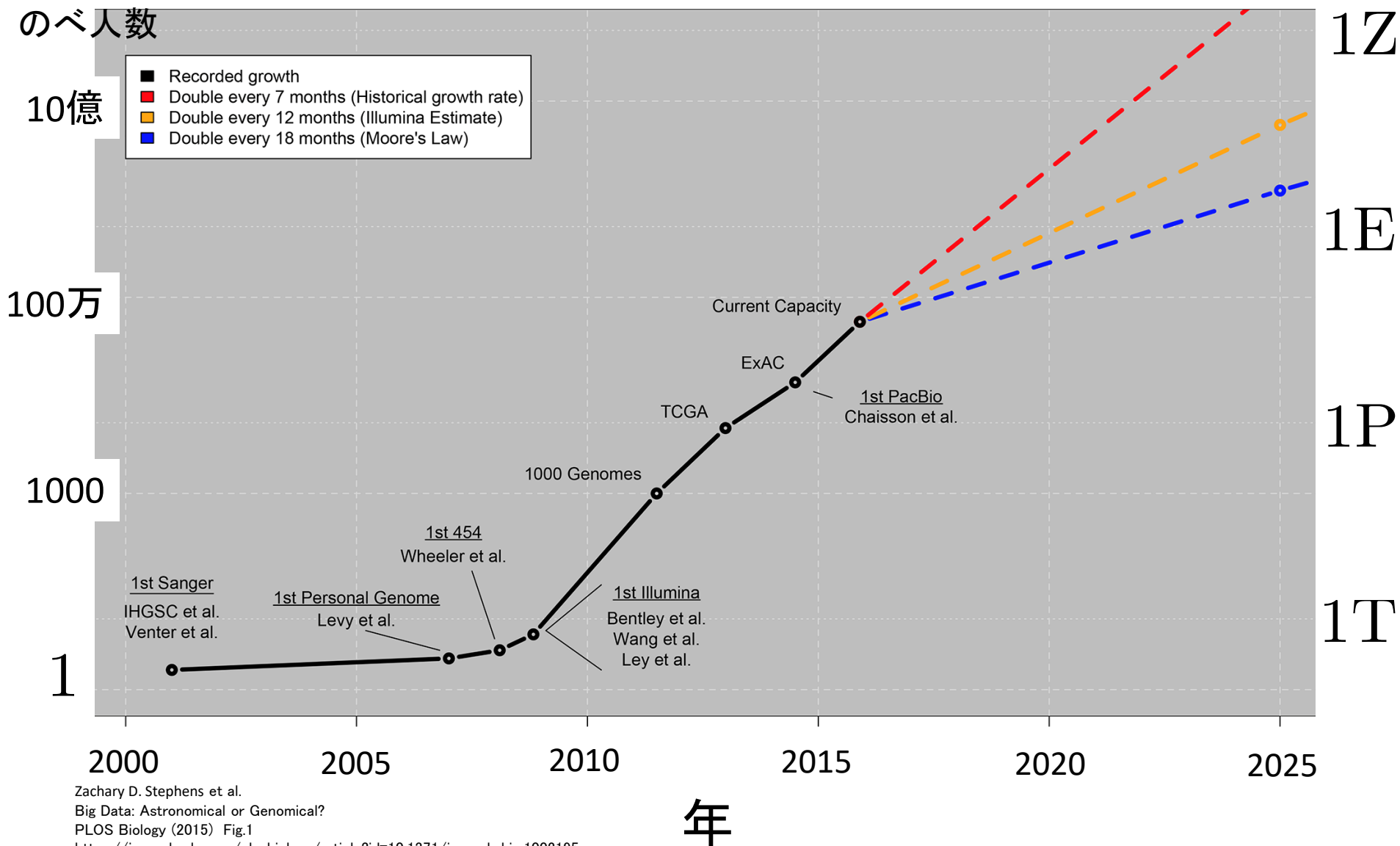
ゲノム情報・環境要因・生活習慣
と
健康・疾患

健康維持・発症予防へ

シーケンスデータの爆発的増加

シーケンスされる
のべ人数

データサイズ [bp]



Zachary D. Stephens et al.
Big Data: Astronomical or Genomical?
PLOS Biology (2015) Fig.1
<https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1002195>
CC BY 4.0



シークエンセータって、 どんなデータ？



生のサンプル

DNAとして抽出
ATCGの4種の文字



イルミナ株式会社
HiSeq X Ten



©2012 Illumina, Inc. All rights reserved.

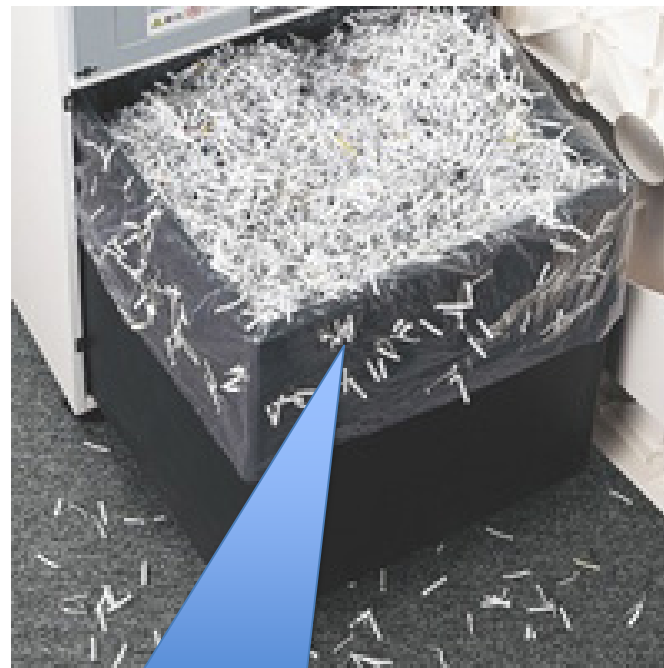
次世代シークエンサー

得られるデータ ゲノムシュレッター

100文字ぐらいの断片になった

12億ピース

の文字列断片がコンピュータに
吐き出される



ATCCGGTAAAT.....TTCA

← 100~150塩基 →

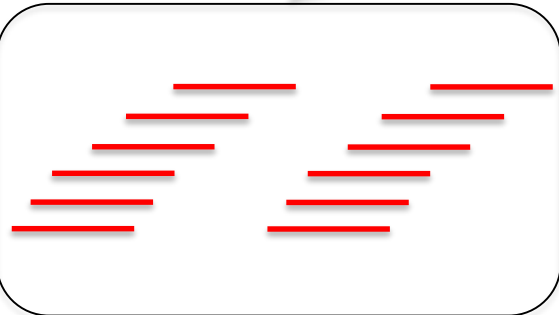


Fastq File

Read(リード)と呼ばれる

```
@HWUSI-EAS1765_0024_FC:8:1:4619:1127#0/1
CTTTATTCCACCTGGTAGGGCATACCCAAGAACCATATACTGAGTCCTGTCTCAGGTTGATGGAGGGTTCCTGGGGCCAAGGCACACAACCTGCCTGTGCTCTGCTC
+HWUSI-EAS1765_0024_FC:8:1:4619:1127#0/1
ggggggggggggffggggcggggggggff_cffffgggggggggggggggffaffededccd¥] ddddggeebf_adfcadc[ad¥aabba]_aaaaaaaa¥O
@HWUSI-EAS1765_0024_FC:8:1:4649:1123#0/1
CGCGCGACTTCCAGGCGGCGCTACGACCGCCACGCGCGGGCCCTGTCCCGAATGCGTCCTCCCCGCCCCCGGTCTCGCCGCGCTCCTCCTGCCCCGTGCGG
+HWUSI-EAS1765_0024_FC:8:1:4649:1123#0/1
ffcffffcW_fffffafcff^aa¥a^K^bb`bb^OO^ZW] ]M]^VYYSS^M^BBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBB
@HWUSI-EAS1765_0024_FC:8:1:4738:1123#0/1
TTCGACTGAGCCGGGATGCCGTGAAGGACTTCGACTGCTGTTGTCTCTCCCTGCAGCCTTGCCACGATCCTGTTGTACCCCAGATGGCGACCTGTATGAGCGTGAG
+HWUSI-EAS1765_0024_FC:8:1:4738:1123#0/1
¥_ZQX^acccddcY_dff_ccfdd_dccffggg¥gSffdcaeWcdff_f¥ffSadfffdWcaff]_cffffd_dfddfd]ccWQMVYIZ¥VZ^WW^__WWacV_
@HWUSI-EAS1765_0024_FC:8:1:4782:1118#0/1
AATGTCTCCGAAGATTTTCATGTGTTTGCAAGATCGCCAGCCTTACAGGTTTTACGTGAAATGAAAGCCATTGGAATAGAACCCTCGCTTGCAACATATCACCATATT
+HWUSI-EAS1765_0024_FC:8:1:4782:1118#0/1
adfagggegggggggggggggfff]ff^fdfggggggefffgfgcdfddfcffffdeggeggggfgggcgdggggge^b]dWfacacdddbcgfefffdb]Vb]```]ea
@HWUSI-EAS1765_0024_FC:8:1:5131:1119#0/1
CTTCGATATGGGATAATCCTATTTATTACCTCAGAAAGTTTTTTTCTTCGCAGGATTTTTCTGAGCCTTTTACCCTCCAGCCTAGCCCCCTACCCCCCACTAGGAGG
+HWUSI-EAS1765_0024_FC:8:1:5131:1119#0/1
gfggggggcgcgg_caffffdgfggggegfggggggffddegggg`egggggggggggfgfggggdgggfggfggfeggcfc]gabbaaafbcdccb`]c_aPMaBB
@HWUSI-EAS1765_0024_FC:8:1:5306:1127#0/1
GCAGGTTTCCCGTGTTCGCAGCGGAGCCGAGGCCAGCTGAACCCGGCCGTGGGAGCACGGCTCGGAGAGAGCGGAAGCGCGGTTCGCAGGAAAGCGGCGACCGCT
+HWUSI-EAS1765_0024_FC:8:1:5306:1127#0/1
ddddcbfff_bdWdbfffcababbdccfc`fcW]YYJQW```]`ae_]_BBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBB
@HWUSI-EAS1765_0024_FC:8:1:5429:1121#0/1
GAAAGAAAGAAAGAGTGGGTTTTCAGAGTGTCAGGGTGGAGTGATTCCAAGTGGGATGGTTGTGGCATTAGTTTAAGCTGAATAAATAATTTCAATTTGGGGCAGTT
+HWUSI-EAS1765_0024_FC:8:1:5429:1121#0/1
ffdaYffcfccfffcfffcfffffddcfdffffcfafe[ffc`cf]fbaXbdb`Wccacb[dc]b]``^`]c`cdaadd`be`]bdbdbaadb_]`T_R¥[
@HWUSI-EAS1765_0024_FC:8:1:5572:1125#0/1
TTGAAATCTGGGAGTTGGTGGGGCAAGGTCACCTCTTGGCAGCTTCTTCTTGGGCAGCCAAATCTGCCTCCTTCTGAGCAGCCAGGAAGATGGCTCGTTCCTTCTG
+HWUSI-EAS1765_0024_FC:8:1:5572:1125#0/1
cedcaYdffffgggggfafe[]ff_dffdaWfffffcfcff]fafd[acafdefcdggggcgddadR]bb_abdabc]a_cddba[]X_]WYZZZN¥YT[¥__[_]
@HWUSI-EAS1765_0024_FC:8:1:5593:1128#0/1
CGATCCTATGTGGAATGTCCGCCACCTTGTTCCAGTGCTGGAAAACGATATTACCTCAGTTTGTAAGGCTGCCAAGTTACTCCATTGGAATGGACATTTGAAGC
```

GACCC**C**TC...
ACATGGACCC**C**TC...
CTCACATGGACCC**C**TC...
TCTGACATGGACCC**A**TC...
...AAGTCTCACATGGACCC**C**TC...
...AAGTCTGACATGGACCC**C**TC...
...**AAGTCTGACATGGACCCCTC**...



参照配列

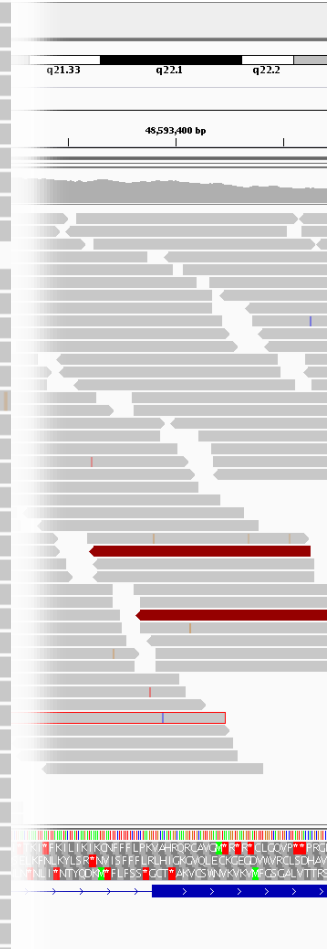
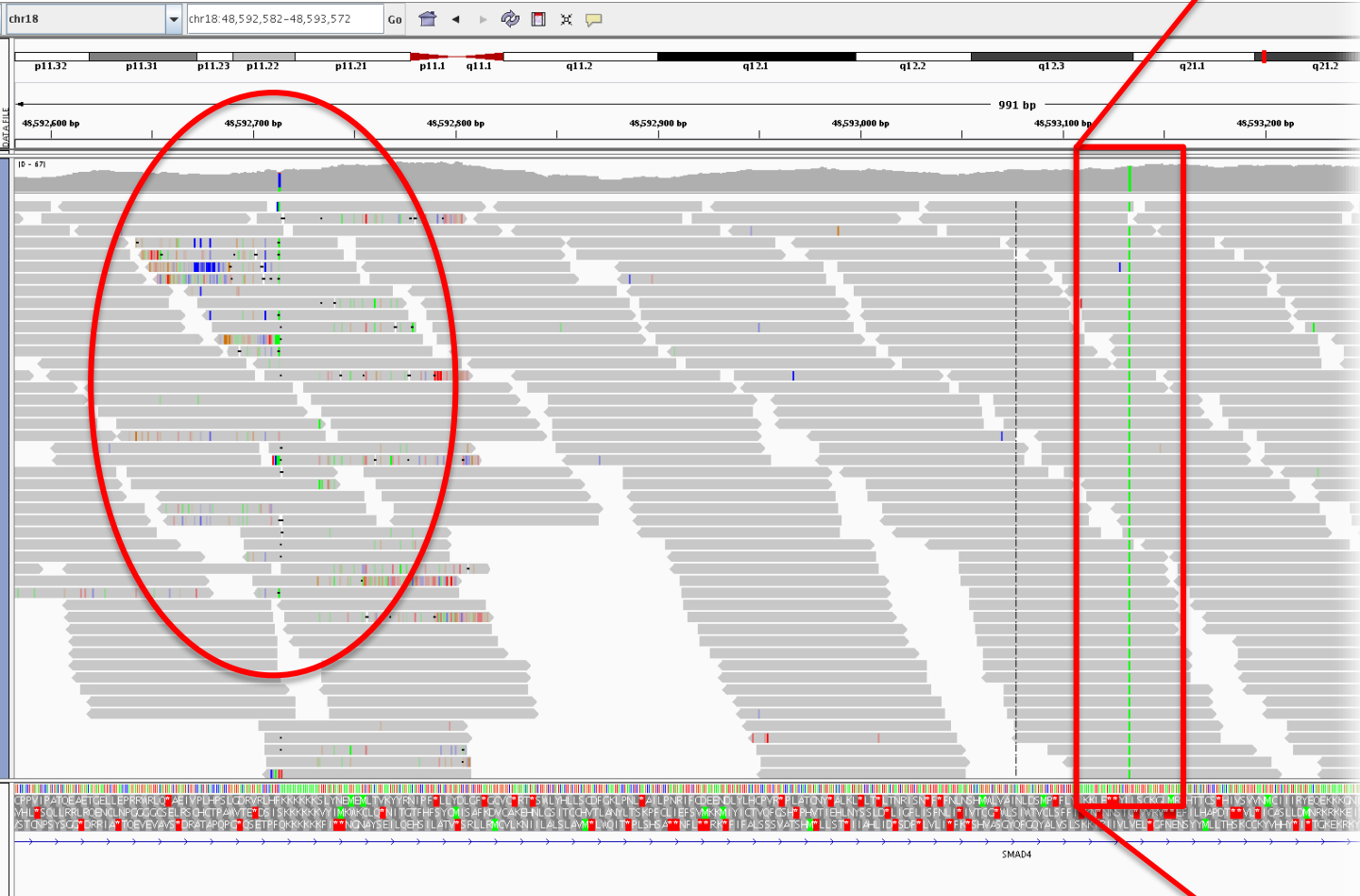
ATGC

IGV Screen Shot



ATGC

SNP



がん

- 2016年にがんで死亡した人は372,986人（男性219,785人、女性153,201人）。（がん統計2018年09月15日より）
- 2014年に新たに診断されたがん（罹患全国合計値）は867,408例（男性501,527例、女性365,881例）。（がん統計2018年09月15日より）
- 人生80年。2人に1人はがんになる。
- ゲノムに生じた変異は、がんの原因となり得る。

ゲノムの変異、色々

元々のゲノム配列 ... **AAGCTCGAATT** ...

一塩基置換 ... **AAGCT****A****GAATT** ...

一塩基欠損 ... **AAGCT****-****GAATT** ...

一塩基挿入 ... **AAGCTC****C****GAATT** ...

フィッシャーの正確確率検定

Yoshida *et al.* (2011) Nature など多数

各位置において、変異を保持するリードの本数を数える

腫瘍組織

ATGCATCAATGCATG
CATGCATCAATGCAT
GCATGCATGAATGCA
TGCATGCATCAATGC
ATGCATGCATCAATG
CATGCATGCATCAAT



	変異あり	変異なし
腫瘍	5	1
正常	0	6

正常組織

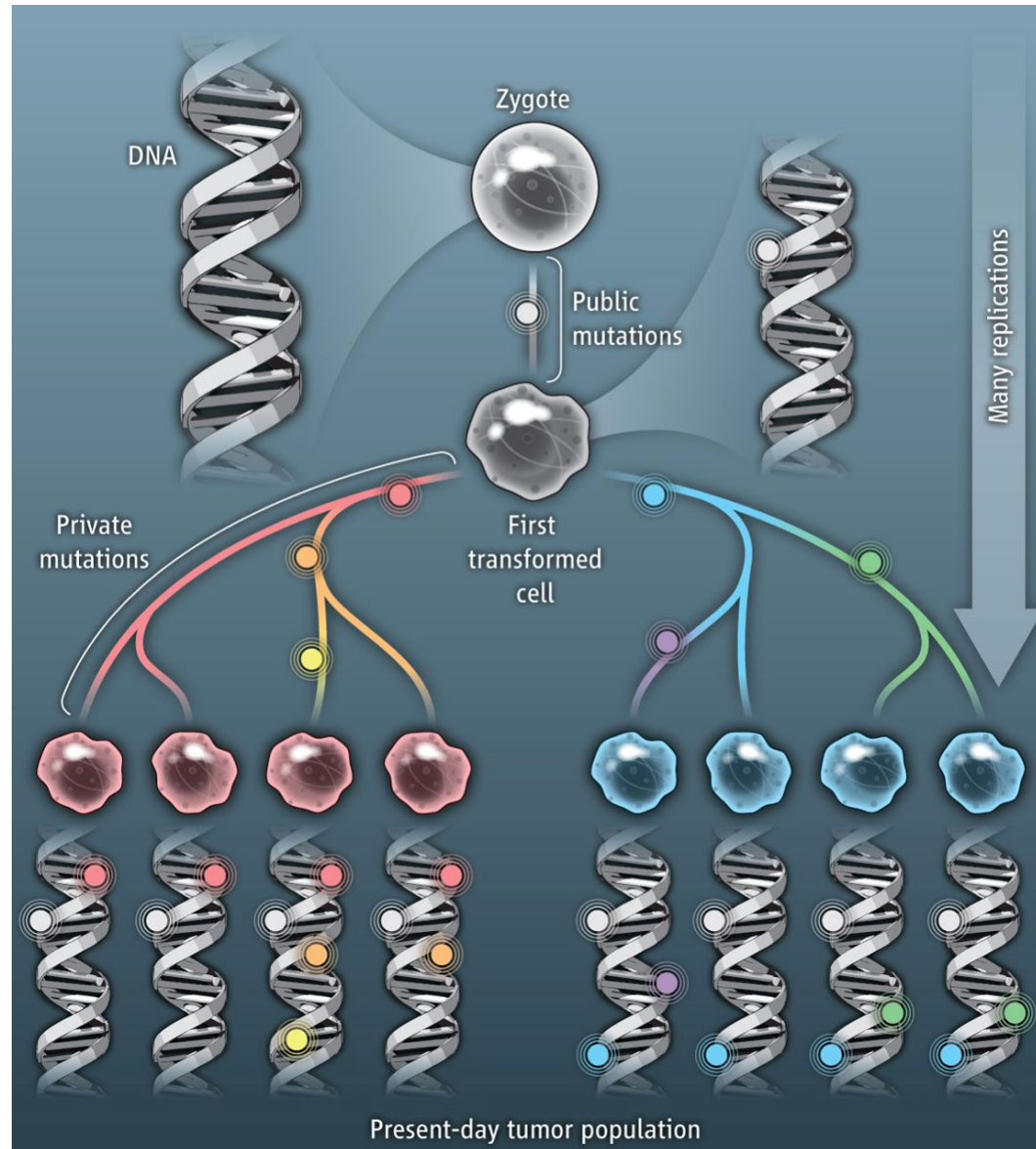
ATGCATGAATGCATG
CATGCATGAATGCAT
GCATGCATGAATGCA
TGCATGCATGAATGC
ATGCATGCATGAATG
CATGCATGCATGAAT



フィッシャーの
正確確率検定

$p\text{-value} = 0.01515$

Cancer Heterogeneity (がんの多様性)



Darryl Shibata
Heterogeneity and Tumor History
Science 20 April 2012: Vol. 336 no. 6079 pp.
304–305, Mutation history. AAAS
<https://science.sciencemag.org/content/336/6079/304.full>



ヒトゲノムシーケンス解析



パネルシーケンス(ターゲットシーケンス) 特定の(遺伝子)領域だけを読み取る 例: 100程度の遺伝子上の変異を確認	2~4万円
全エクソームシーケンス ヒト約2万遺伝子の領域のみを読み取る (全ゲノムの3%程度の領域)	5~10万円
全ゲノムシーケンス 30 億塩基対の全てを読み取る	5~10万円

東大医科研 HGC スーパーコンピュータ SHIROKANE

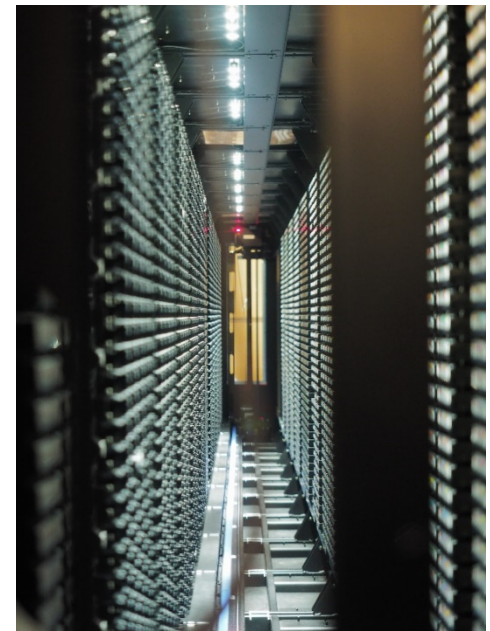
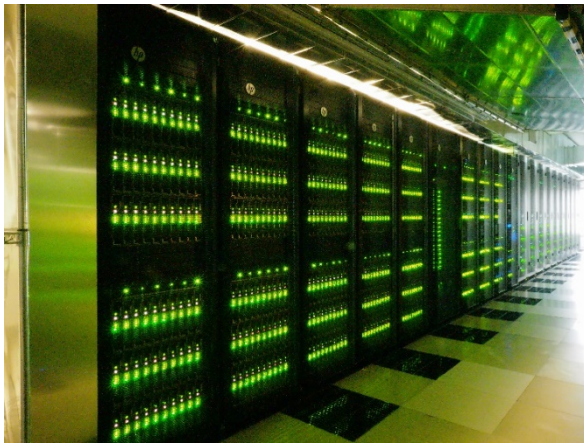
生命・医学系専用の運用をしている
国内最大規模のスパコンです

巨大ストレージ

1.5 Peta FLOPS

18PB高速ディスクアレイ

100PBアーカイブ



<http://www.hgc.jp/~ayumu/705/P4207532.jpg>

Shirokane を、主力の計算機「計算
ノード Thin」側から見たもの。

Shirokane4 と Shirokane5の理論演算
性能は計 1.5 Peta FLOPS
GPU v100 x 80 を搭載
約2万の CPUコア

<http://www.hgc.jp/~ayumu/705/P4207534.jpg>

Shirokane を、「共有ディスク」側から見たもの。全ての計算ノードから同時に
高速にアクセスできるディスクが
18 ペタバイトある

IBM社製テープアーカイブ
+ 3 PB ニアラインディスク

©ヒトゲノム解析センター
<http://hgc.jp/japanese/>

18



Genomon 2



- シーケンスデータから統計学的方法で
- 高感度、高正答率を達成する変異検出
 - Genomon Mutation
 - Genomon-SV
- すぐに利用できる
 - HGCのスパコンにインストールされている
 - 他のスパコンにもインストール可能
 - 特別な設定やソフトウェアインストールなど必要なし
 - 入力データを準備し、一つコマンドを入力するだけ
 - `>> genomon_pipeline dna input.csv output_dir`

Developers: Dr. Yuichi Shiraishi, Kenichi Chiba, Ai Okada et al.



PCA WG (Pan-Cancer Analysis of Whole Genomes)

Challenges

- ✓ TCGA/ICGCのゲノム解析の多くはエキソーム
- ✓ TCGA/ICGCで現在解析している**2,834症例**のがんの**WGSをData sharing** 》日本から肝臓300例(10%)
- ✓ 10 のデータセンターを連携したクラウドを構築
同じ環境化(VM)でのデータ解析を行う(1PB)
》 **東大医科研 Shirokane が参画**
- ✓ 解釈を行う下流解析(16 のテーマ)を国際連携グループで実行
- ✓ **東大医科研は、免疫ゲノム解析拠点を担った**
- ✓ 600人以上の世界中の研究者(含補助)の共同作業



PCA WG
PanCancer Analysis
OF WHOLE GENOMES

2,834 cancer WGS

- Under the same VM environment
- 1. Sanger pipeline
 - 2. DKFZ/EMBL pipeline
 - 3. Broad pipeline



Amazon clouds

Synchronized 10 Cloud Data Centers

クレイ・ジャパン・インク
Appro Xtreme-X



Shirokane

1 PB Chicago

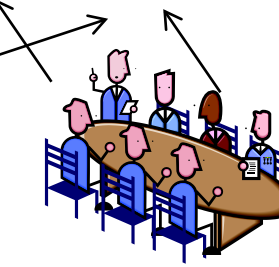
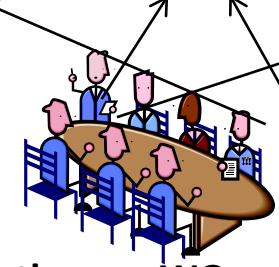
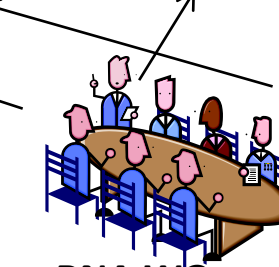
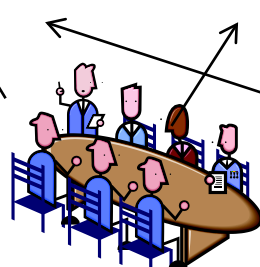
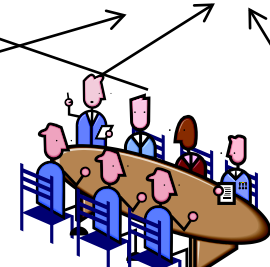
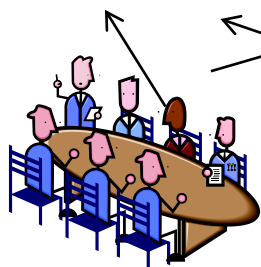
1 PB London

1 PB Heidelberg

1 PB Barcelona

1 PB Seoul

1 PB Tokyo



Signature WG

SV WG

Cis-regulatory WG

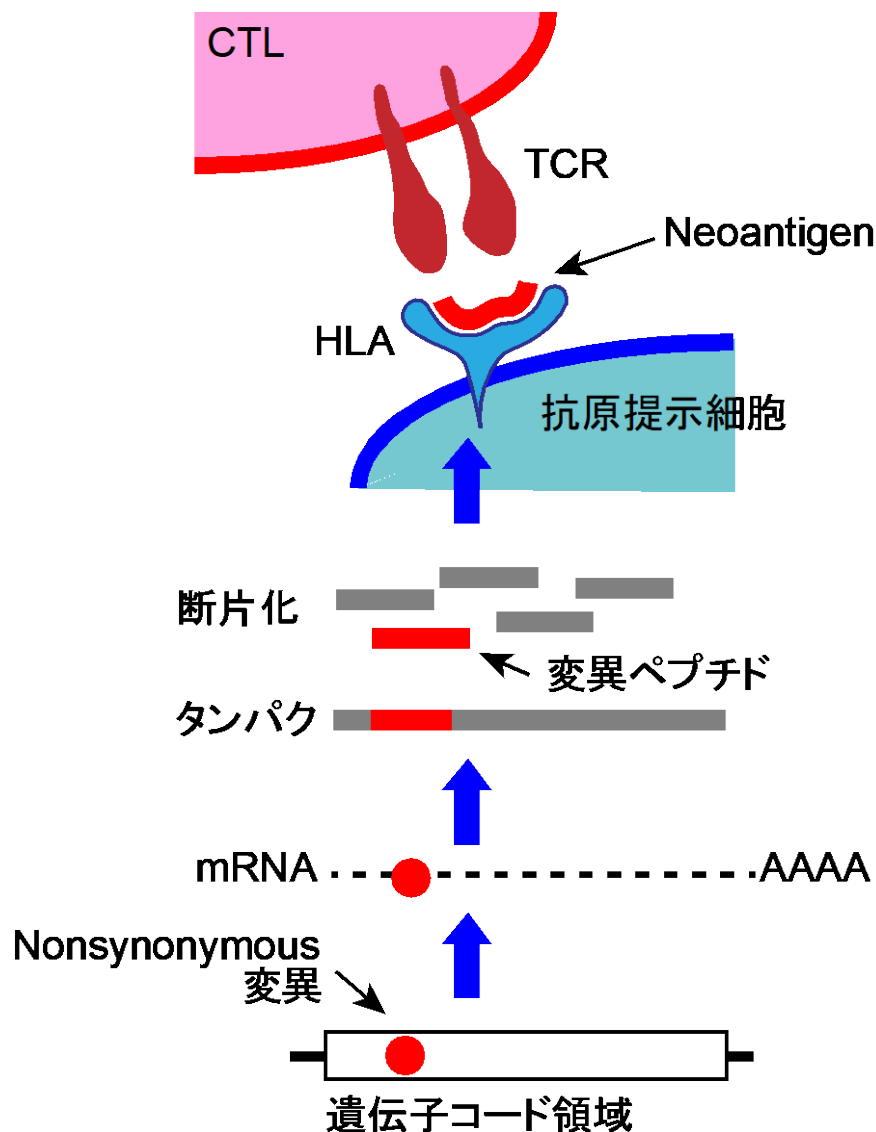
ncRNA WG

Pathogen WG

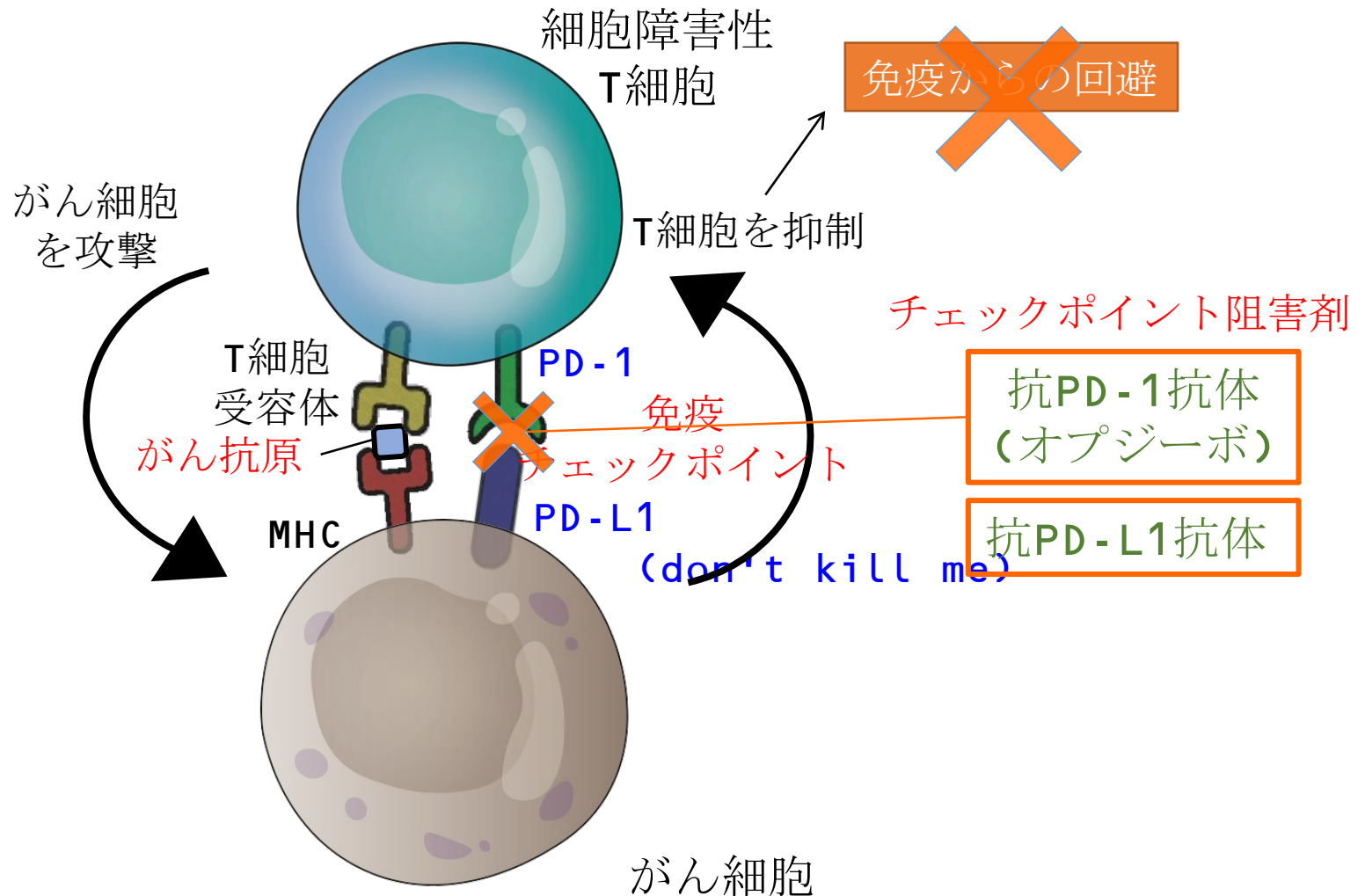
ImmunoGenomics WG

16 working groups and 600 researchers and tech staff world-wide

がんに対する免疫監視システム



生体にはがんを防ぐ免疫のしくみがある



HLA 遺伝子解析

Types of HLA-A gene

HLA-A*02:01:01:01

HLA-A*03:01:01:01

HLA-A*11:01:01:01

HLA-A*24:02:01:01

HLA-A*26:01:01

DNA sequences of HLA-A gene

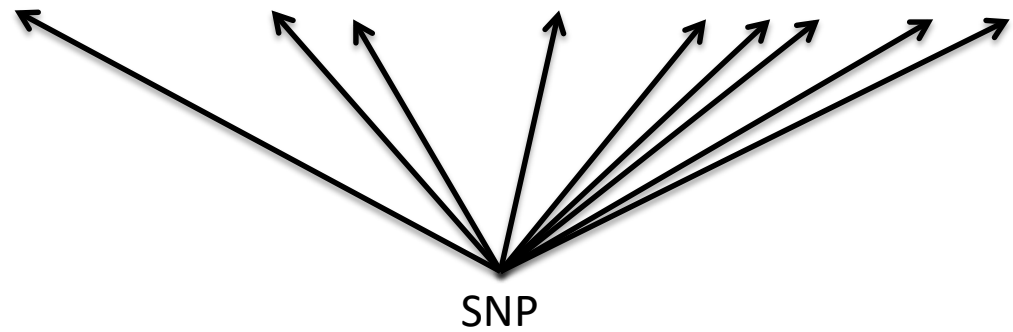
...ACTCACCGAGTGGACCTGGGGACCCCTGCGCGGC...

...ACTGACCGAGTGGACCTGGGGACCCCTGCGCGGC...

...ACTGACCGAGTGGACCTGGGGACCCCTGCGCGGC...

...ACTGACCGAGAGAACCTGCGGATC GCGCTCCGC...

...ACTGACCGAGCGAACCTGGGGACCCCTGCGCGGC...



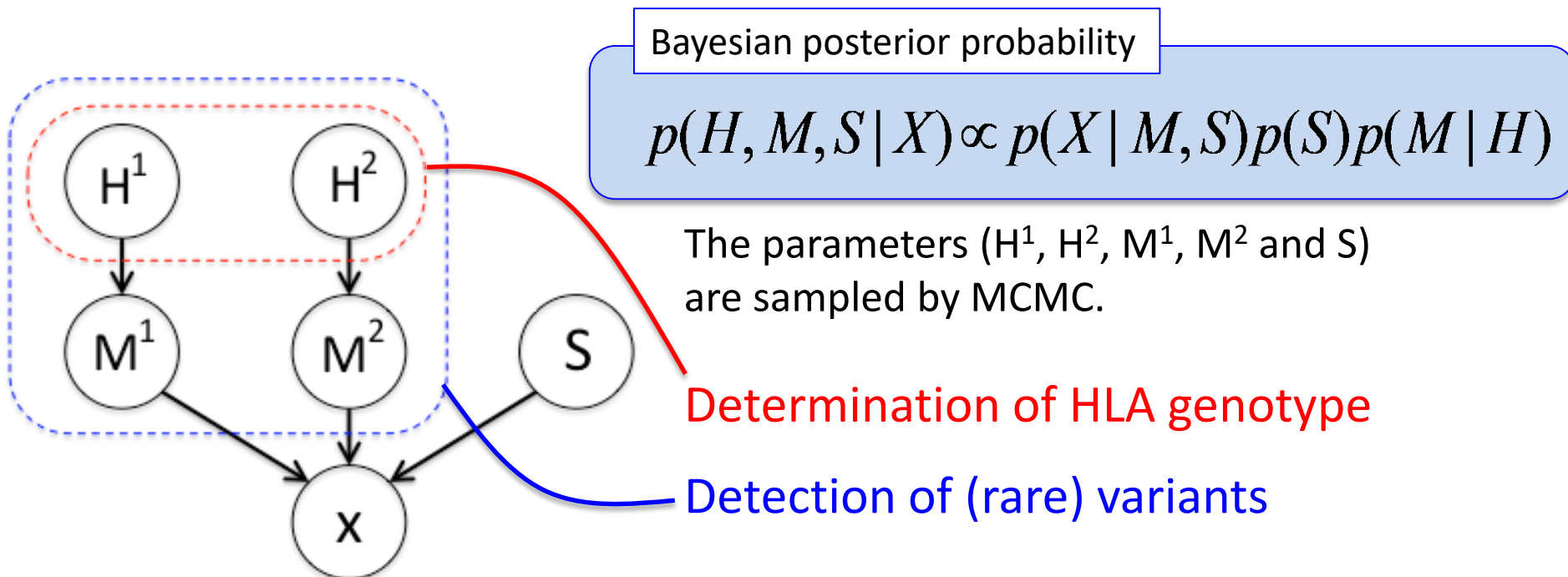
HLA 遺伝子型決定のベイズモデル

H^1, H^2 : HLA 型候補 (参照配列)

M^1, M^2 : サンプルの HLA 配列 (データからの予測)

S : リード x が生成された HLA 型 (M^1 or M^2) を指示する
ダミー変数

x : ペアエンドリード (シングルエンドのリードでも可)



Performance Evaluation: HLA genotyping on Japanese 25 WGSs

ALPHLARD: a Bayesian method for analyzing HLA genes
from whole genome sequence data, Table 1
Shuto Hayashi, Seiya Imoto, et al.
bioRxiv 2018
<https://www.biorxiv.org/content/10.1101/323766v1.full.pdf>

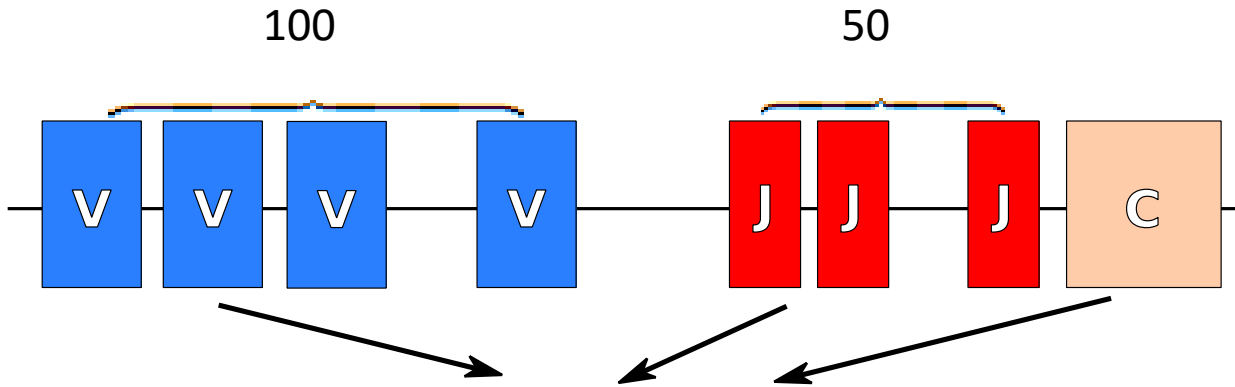
		ALPHLARD	OptiType	PHLAT	HLA-VBSeq
HLA-A	2-digit	100% (50/50)	100% (50/50)	76.0% (38/50)	96.0% (48/50)
	4-digit	98.0% (49/50)	98.0% (49/50)	60.0% (30/50)	82.0% (41/50)
	6-digit	98.0% (49/50)	N/A	46.0% (23/50)	82.0% (41/50)
HLA-B	2-digit	100% (48/48)	87.5% (42/48)	72.9% (35/48)	89.6% (43/48)
	4-digit	100% (48/48)	85.4% (41/48)	56.3% (27/48)	75.0% (36/48)
	6-digit	95.8% (46/48)	N/A	39.6% (19/48)	72.9% (35/48)
HLA-C	2-digit	100% (50/50)	100% (50/50)	78.0% (39/50)	96.0% (48/50)
	4-digit	98.0% (49/50)	94.0% (47/50)	56.0% (28/50)	66.0% (33/50)
	6-digit	98.0% (49/50)	N/A	44.0% (22/50)	66.0% (33/50)
HLA-DPA1	2-digit	100% (24/24)	N/A	N/A	87.5% (21/24)
	4-digit	100% (24/24)	N/A	N/A	87.5% (21/24)
	6-digit	100% (24/24)	N/A	N/A	87.5% (21/24)
HLA-DPB1	2-digit	100% (22/22)	N/A	N/A	86.4% (19/22)
	4-digit	100% (22/22)	N/A	N/A	86.4% (19/22)
	6-digit	100% (22/22)	N/A	N/A	86.4% (19/22)
HLA-DQA1	2-digit	100% (24/24)	N/A	70.8% (17/24)	100% (24/24)
	4-digit	95.8% (23/24)	N/A	62.5% (15/24)	95.8% (23/24)
	6-digit	95.8% (23/24)	N/A	62.5% (15/24)	95.8% (23/24)
HLA-DQB1	2-digit	100% (18/18)	N/A	77.8% (14/18)	100% (18/18)
	4-digit	94.4% (17/18)	N/A	61.1% (11/18)	88.9% (16/18)
	6-digit	94.4% (17/18)	N/A	38.9% (7/18)	88.9% (16/18)
HLA-DRB1	2-digit	100% (24/24)	N/A	70.8% (17/24)	95.8% (23/24)
	4-digit	100% (24/24)	N/A	50.0% (12/24)	58.3% (14/24)
	6-digit	100% (24/24)	N/A	45.8% (11/24)	58.3% (14/24)
Total	2-digit	100% (260/260)	95.9% (142/148)	74.8% (160/214)	93.8% (244/260)
	4-digit	98.5% (256/260)	92.6% (137/148)	57.5% (123/214)	78.1% (203/260)
	6-digit	97.7% (254/260)	N/A	45.3% (97/214)	77.7% (202/260)

Collaboration with
Dr. Hidewaki Nakagawa
at Riken

T-cell レパトア

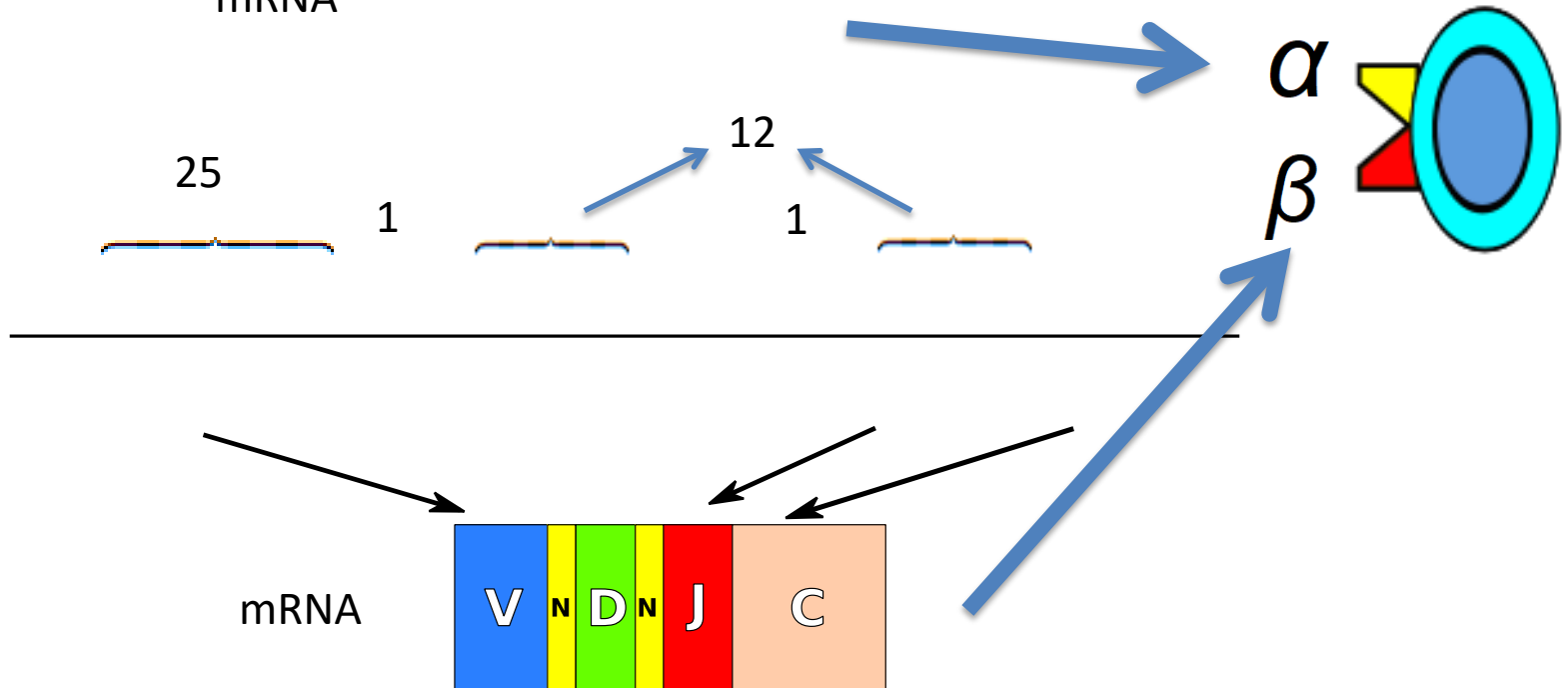
Acquisition of Diversity through Genomic Rearrangement

α chain
(Chr 14)



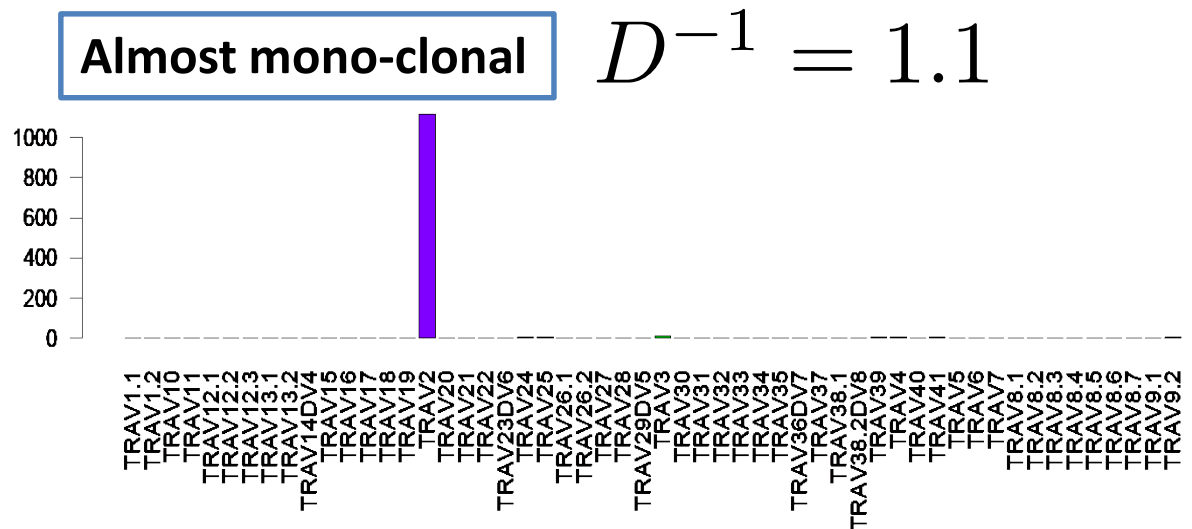
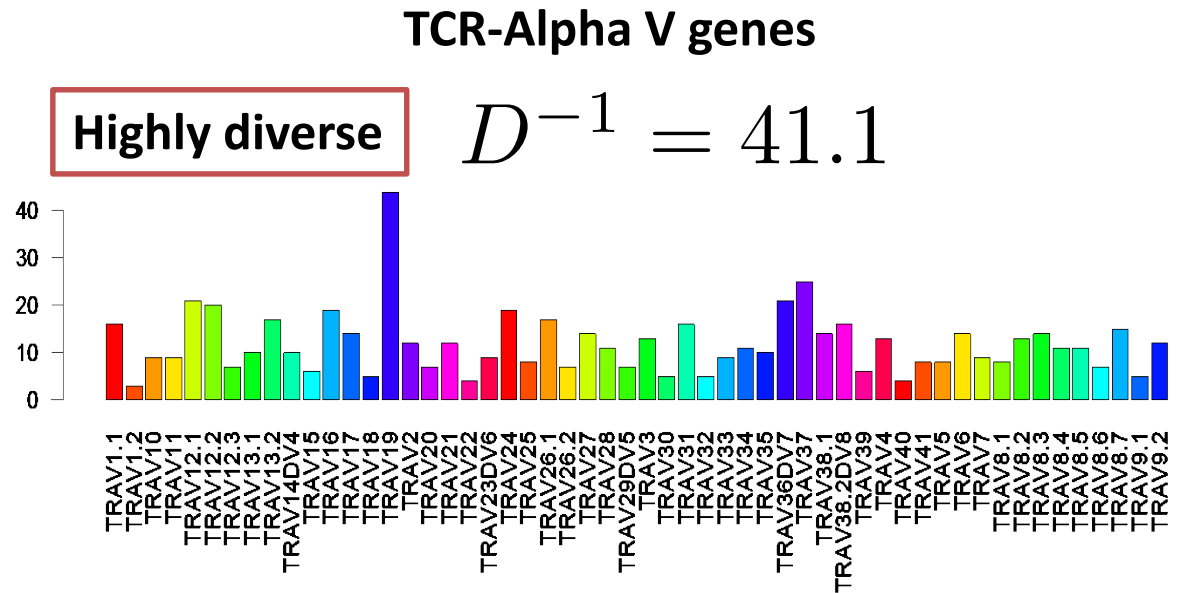
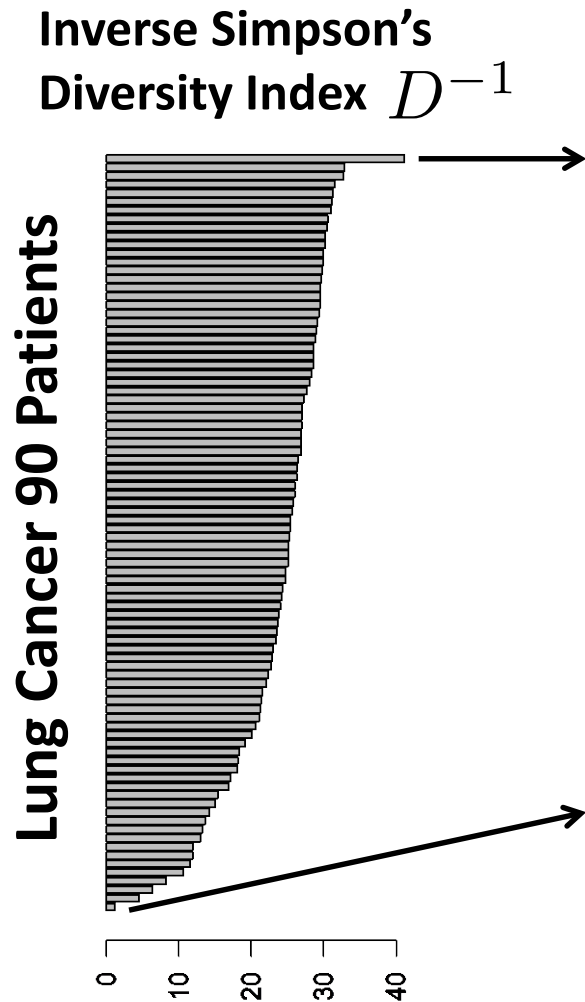
mRNA

β chain
(Chr 7)



mRNA

RNA-seq: TCR Repertoire Analysis



スパコン使えます

The screenshot shows the Human Genome Center website. At the top, there's a navigation bar with links to 'サイトマップ' and 'English'. Below that is the Human Genome Center logo and name, along with a search bar. A banner for 'スパコン' (Supercomputer) is visible. The main content area is titled 'スーパーコンピュータ' (Supercomputer) and includes sections for '民間利用・有償利用サービス' (Private/For-profit service), '利用資格・利用方法' (Eligibility and Usage), and 'HGC のスパコンを利用するメリット' (Benefits of using HGC's supercomputer). A sidebar on the right lists various resources like 'SHIROKANE 利用案内' (SHIROKANE User Guide) and '利用方法、講習会' (Usage and Workshops).

Human Genome Center
the Institute of Medical Science, the University of Tokyo

企業の方もスパコンを利用できます
利用料金の詳細

Google カスタム検索

トップページ 解析ツール データベース Anonymous スーパーコンピュータ HGC とは

TOP > スーパーコンピュータ > 利用案内 民間利用・有償利用サービス FTP

スーパーコンピュータ

民間利用・有償利用サービス

利用案内

当センターのスパコンシステムは、学術機関だけでなく、民間機関もご利用いただけます。当センターは個別に調達、維持することが困難なスパコンを安価に民間機関に提供することで、ゲノム医学分野の発展を支えます。

利用資格・利用方法

● 利用資格

民間利用・有償利用サービスでは、[こちらの](#)ページの「申請資格」のいずれかに該当する方がスパコンを利用できます。

● 利用方法

[こちらのページの案内](#)の新規利用申請をご提出ください。受理の後、審査を行ない、アカウント発行いたします。通常、5 日程度でアカウントを発行しています。

HGC のスパコンを利用するメリット

当センターのスパコンシステムは、単純にシステムを提供するホスティングサービスを行っている訳ではありません。当センターのスパコンシステムには以下のメリットがあります。

- 購入するよりも安い価格で多くの計算機資源を利用できる。
- スパコンならではの[高速な共有ディスク](#) (Lustre ファイルシステム) を備え、一般的サーバよりも楽に、高速なファイルアクセスができる。
- GenBank RefSeq UniProt 等の公共データベースのデータのミラーが存在している。

SHIROKANE 利用案内

- [スパコン利用案内](#)
- [民間利用サービス](#)
- [利用料金](#)
- [新規利用手続き](#)
- [継続利用手続き](#)
- [各種申請手続き](#)
- [技術支援サービス](#)
- [情報セキュリティポリシー](#)
- [よくあるお問い合わせ](#)

利用方法、講習会

- [利用上の注意](#)
- [講習会日程](#)
- [講習会資料](#)
- [スパコン Wiki](#)
- [SHIROKANE Q&A](#)
- [ユーザマニュアル](#)
- [Linux 関連利用法](#)
- [ミラーデータベース利用法](#)
- [Univa Grid Engine 利用法](#)
- [ハードディスククラッシュ](#)
- [スパコン利用の参考文献](#)
- [バックアップポリシー](#)

電気代程度の料金を頂いています。

(例)
年間36万円で
・6 TB の高速ディスク
・32 TB のデータ保存用
アーカイブ
・1,284コア CPU
が使えます。(学術利用)



全ゲノム解析の時代に

著作権等の都合により省略しました

genomewebウェブサイトの画像

Whole-Genome Sequencing Has Better Dx Rate Than
Exome at Similar Cost, Australian Team Finds

2016年5月18日

<https://www.genomeweb.com/sequencing/whole-genome-sequencing-has-better-dx-rate-exome-similar-cost-australian-team-finds#.XQbysoj7SUK>



研究では、全ゲノムシーケンス 当たり前だが



全ゲノムシーケンスデータ = 200GB 程度

ヒトゲノム標準配列と異なる部分 = 400万～500万箇所

がん細胞の変異 = 数千～数万箇所

臨床シーケンスへの実装



ビッグデータを個人に還元する取り組み 臨床シークエンスとは？



ゲノム

一人あたり数GBから数百GBのデータ

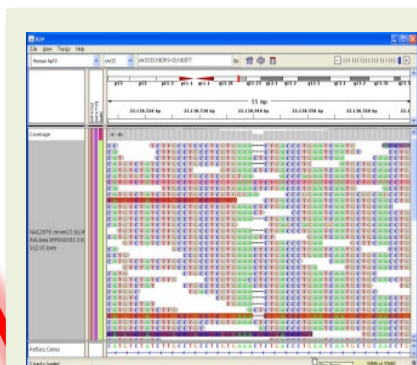
数百から数百万のDNA配列の違い



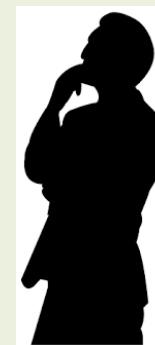
患者さん



主治医



ゲノム解析
DNA



解釈・翻訳

個人ゲノムに基づく個別化医療



Nick Volkar君



Photos by Gary Porter

著作権等の都合により
省略しました

By Lou Saldivar and Gary Porter
Nicholas' DNA: Unraveling the
mystery

December 21, 2010

MSOE Center for BioMolecular
Modeling

[http://cbm.msOE.edu/markMyweb/
ggpmResources/documents/
xiapMystery.pdf](http://cbm.msOE.edu/markMyweb/ggpmResources/documents/xiapMystery.pdf)

原因不明の
腸の炎症に苦しんでいた
5歳の男児

SequencerでDNAを解読
(Exome):
16,124変異からの絞り込み

結果XIAPというX染色体上の
免疫にかかわる遺伝子に変異
が見つかり、臍帯血移植により
寛解

臨床シーケンス

世界で最初の パーソナルゲノム医療

2007年5月、ウィスコンシン小児病院
に2歳の男の子がやって来た。

食事をするたびに腸に小さな穴が開き、
その穴が皮膚表面まで通じてそこから
便が漏れるという奇病を患っている。

2009年、臨床の間では世界に例のない
ゲノム解析により、原因遺伝子を突
きとめて治療の手がかりをつかもうと
いう大胆な試みに踏みきる。

診断名のつかない難病を抱えた少年
との出会いから、世界初のパーソナル
ゲノム医療が実現するまでを息詰まる
筆致で綴った医療ドキュメンタリー。
2011年「ピューリッツァー賞・解説報道
部門」受賞記事の書籍化！

マーク・ジョンソン(著/文)キャスリーン・ギャラガー(著/文)梶山 あゆみ(翻訳)井元 清哉(解説)
『10億分の1を乗り越えた少年と科学者たち——世界初のパーソナルゲノム医療はこうして実
現した』(紀伊国屋書店出版部、2018年)

2018年11月1日発売

UTokyo Online Education 学術俯瞰講義 2019 井元清哉 CC BY-NC-ND

マーク・ジョンソン＋
キャスリーン・ギャラガー
梶山あゆみ＝訳
井元清哉＝解説

One in
a Billion

The Story of Nic Volker
and the Dawn of
Genomic Medicine
Mark Johnson &
Kathleen Gallagher

10億分の1を 乗り越えた少年と 科学者たち

世界初のパーソナルゲノム医療はこうして実現した

ピューリッツァー賞
解説報道部門受賞
記事を書籍化！

2007年5月、ものを食べると腸に穴が開き、
皮膚から便が漏れるという奇病を患った

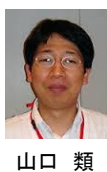
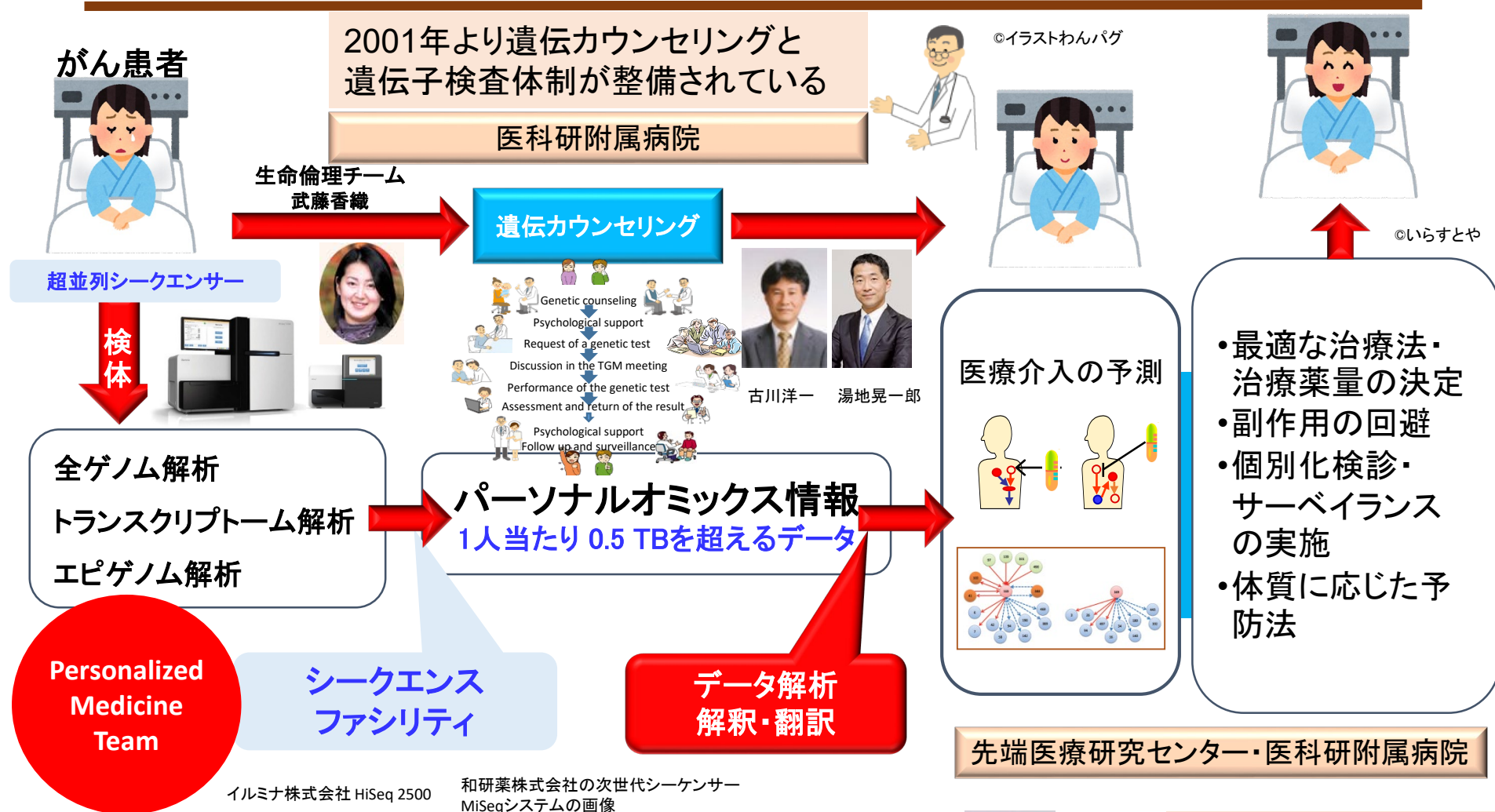
2歳の少年がウィスコンシン小児病院に運ばれる。

“10億人にひとり”レベルの症例で診断名もつかない。

このままでは命がもたないと思われた。万策つきた医師たちは2009年、
最後の手段として臨床の間では世界に例のないゲノム解析により、

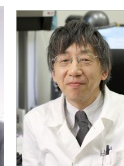
原因遺伝子を突きとめるという大胆な試みに踏みきる——

がん個別化ゲノム医療の実践体制 (東大医科学研究所及び研究所附属病院)



ヒトゲノム解析センター・ヘルスインテリジェンスセンター

スパコン: 国際がんゲノムコンソーシアムを初めとする全ゲノムシーケンシング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



がん臨床シークエンス Team IMSUT



消化器がん

- 古川 洋一
- 池上 恒雄
- 山口 貴世志
- 高根希世子
- 畠山 晴良
- 幸保 莉香
- 斎藤紀美子
- 宮城 洋平

血液がん

- 東條 有伸
- 横山 和明
- 小林 真之
- 中村 聡介
- 磯部 優理
- 武井 智美
- 小川 弥穂
- 近藤幹也
- 山崎 美香
- 磯部友見子
- 安井 寛
- 湯地 晃一郎

インフォマティクス

- 宮野 悟
- 渋谷 哲朗
- 山口 類
- 片山 琴絵
- 長谷川 嵩矩
- 齊藤 あゆむ
- 清水 英悟
- 笠島 理加
- 山本 茉莉
- 上村 光弘
- 油井 宏昭

ELSI

- 武藤 香織
- 井上 悠輔
- 神里 彩子



- **固形腫瘍(大腸がんなど)**
(古川洋一教授)
 - FAP, Lynch, PMP, etc.
 - 全ゲノム解析
- **血液腫瘍(急性骨髄性白血病など)**
(東條有伸教授)
 - MDS, ALL, etc.
 - パネルシーケンス



TruSight Myeloid Panel (Illumina)



ABL1	CEBPA	HRAS	MYD88	SF3B1
ASXL1	CSF3R	IDH1	NOTCH1	SMC1A
ATRX	CUX1	IDH2	NPM1	SMC3
BCOR	DNMT3A	IKZF1	NRAS	SRSF2
BCORL1	ETV6/TEL	JAK2	PDGFRA	STAG2
BRAF	EZH2	JAK3	PHF6	TET2
CALR	FBXW7	KDM6A	PTEN	TP53
CBL	FLT3	KIT	PTPN11	U2AF1
CBLB	GATA1	KRAS	RAD21	WT1
CBLC	GATA2	MLL	RUNX1	ZRSR2
CDKN2A	GNAS	MPL	SETBP1	



Myeloid Panel 臨床シーケンス



がん関連遺伝子のパネルシーケンス

2016年7月

54遺伝子（血液がんとの関連が良く研究されている）

医科学研究所においても、昨年度 250 サンプル以上（血液がん）を実施

診断も薬を選ぶこともできなかった患者さんに対して
パネルシーケンス解析を行った結果

残念ながら・・・
診断にも薬の選択
にも繋がらなかった

53%

全エキソーム解析
全ゲノム解析

27% 診断が付いた

薬の選択に
活用できた
7%

13% 診断と薬の選択
両方に活用



数千行も続く変異情報



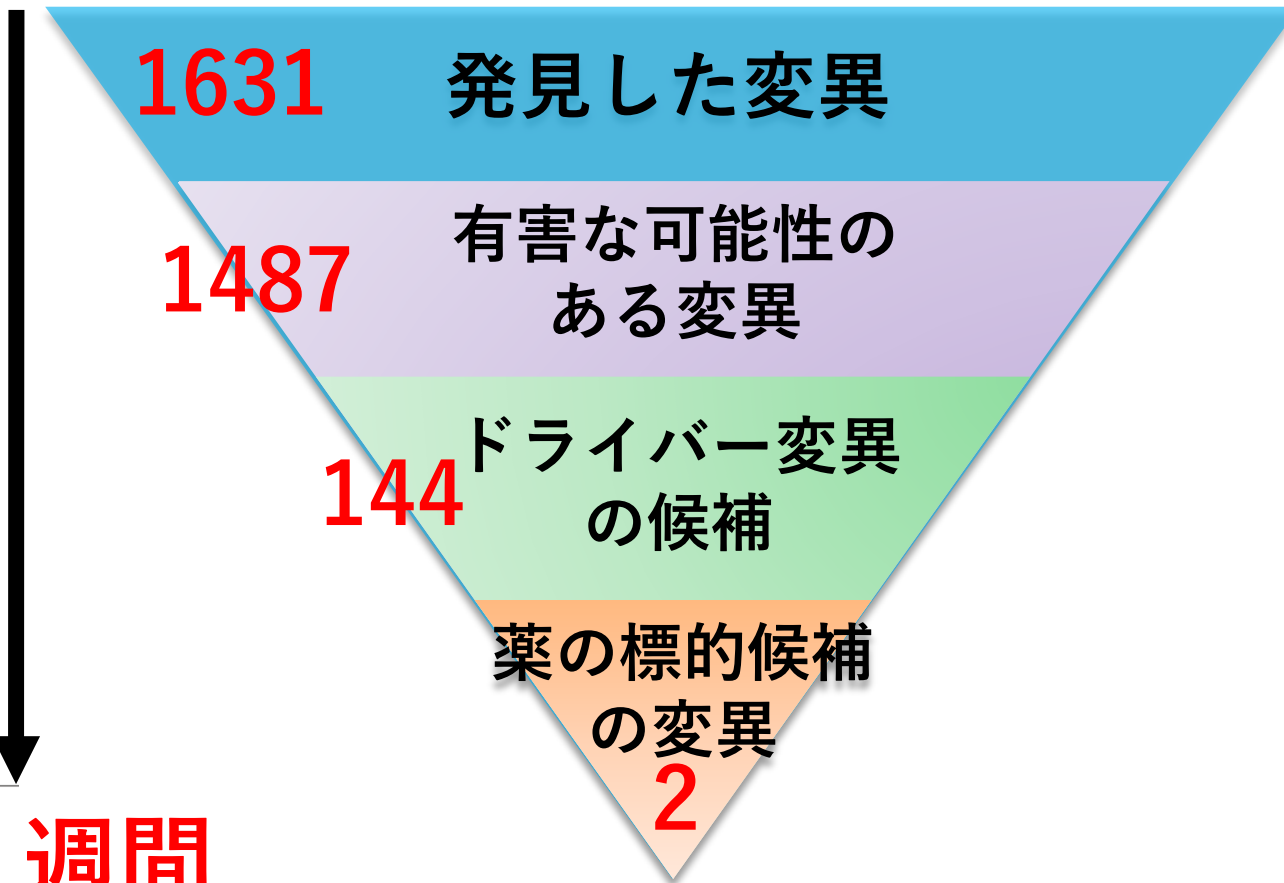
B	C	D	E	F	G	H	I	J	K	L	M	N	O
Funcio	Gene name, transcript identifier and the sequence change in the c	Chr	Start	End	Ref	Obs	bases.tu	bases.n	misRate	strandR	misRate	strandR	p-value
nonsynony	Pex26:NM_028730:exon5:c.G866C:p.R289P,	chr6	1.21E+08	1.21E+08	G	C	0,320,0,0	0,0,552,0	1	0.49375	0	---	247.3795
synonymou	Rp1:NM_011283:exon4:c.T972A:p.I324I,	chr1	4339997	4339997	A	T	58,0,0,51	183,0,0,0	0.46789	0.490196	0	---	25.92997
nonsynony	Atp7b:NM_007511:exon17:c.C3660A:p.N1220K,	chr8	23108727	23108727	G	T	0,0,0,185	0,0,405,0	1	0.405405	0	---	157.9054
synonymou	Prdm14:NM_001081209:exon4:c.A859C:p.R287R,	chr1	13114467	13114467	T	G	0,0,38,28	0,0,0,112	0.575758	0.447368	0	---	20.414
nonsynony	Olfir452:NM_001011869:exon1:c.A788T:p.H263L,	chr6	42740827	42740827	A	T	50,1,0,245	447,0,0,0	0.827703	0.428571	0	---	145.8038
nonsynony	Kcnu1:NM_008432:exon11:c.G1157T:p.C386F,	chr8	26997228	26997228	G	T	0,0,0,228	0,0,260,0	1	0.372807	0	---	145.0051
nonsynony	AF366264:NM_153093:exon1:c.G730T:p.D244Y,	chr8	13837360	13837360	C	A	262,44,0,0	2,390,1,1	0.856209	0.637405	0.005076	0.5	141.1095
frameshift	Olfir742:NM_146430:exon1:c.694delC:p.L232fs,	chr14	51135574	51135574	C	-	203,179	361,0	0.881773	0.52514	0	---	120.643
nonsynony	Ptprh:NM_207270:exon9:c.G1504C:p.G502R,	chr7	4521076	4521076	C	G	0,0,116,0	0,375,0,0	1	0.663793	0	---	115.209
nonsynony	Pcdhb10:NM_053135:exon1:c.A1931C:p.K644T,	chr18	37573457	37573457	A	C	0,128,0,0	311,0,1,0	1	0.429688	0	---	113.6924
nonsynony	Slco1a1:NM_013797:exon7:c.G872T:p.R291I,	chr6	1.42E+08	1.42E+08	C	A	141,0,0,0	0,246,0,0	1	0.567376	0	---	108.8586
nonsynony	Bms1:NM_194339:exon4:c.G400T:p.G134W,	chr6	1.18E+08	1.18E+08	C	A	119,0,0,0	0,297,0,0	1	0.344538	0	---	106.78
nonsynony	Tg:NM_009375:exon38:c.C6592A:p.P2198T,	chr15	66595862	66595862	C	A	204,86,0,0	0,391,0,0	0.703448	0.52451	0	---	103.805
nonsynony	Vmn1r49:NM_011911:exon1:c.G233C:p.G78A,	chr6	90022780	90022780	C	G	0,49,192,0	0,315,0,1	0.79668	0.484375	0	---	102.5208
nonsynony	Krt78:NM_212487:exon9:c.A275G:p.T919A,	chr15	1.02E+08	1.02E+08	T	C	0,214,0,29	0,0,0,786	0.424603	0.551402	0	---	102.4206
nonsynony	Muc5b:NM_028801:exon31:c.G5564T:p.G1855V,	chr7	1.49E+08	1.49E+08	G	T	0,0,328,22	0,0,762,0	0.407942	0.646018	0	---	99.37521
nonsynony	Pax4:NM_001159926:exon6:c.C536A:p.P179Q,Pax4:NM_011038:exon6	chr6	28395139	28395139	G	T	0,0,33,171	0,0,301,1	0.838235	0.251462	0.003311	0	98.86067
synonymou	Il18r1:NM_008365:exon11:c.G1455C:p.S485S,	chr1	40555376	40555376	G	C	0,23,25,0	0,0,79,0	0.479167	0.695652	0	---	11.56289
nonsynony	Gsdmc4:NM_028992:exon9:c.C905A:p.S302Y,	chr15	63725349	63725349	G	T	0,0,60,237	0,0,252,0	0.79798	0.443038	0	---	97.60669
nonsynony	Tpr:NM_133780:exon38:c.C5519A:p.T1840K,	chr1	1.52E+08	1.52E+08	C	A	110,0,0,0	0,265,0,0	1	0.454545	0	---	97.20361
nonsynony	Rrad:NM_019662:exon2:c.G364A:p.A122T,	chr8	1.07E+08	1.07E+08	C	T	0,35,0,153	0,307,0,0	0.81383	0.522876	0	---	93.4094
nonsynony	Bod1l:NM_001081422:exon10:c.G4950T:p.E1650D,	chr5	42210259	42210259	C	A	148,59,0,0	0,390,0,0	0.714976	0.398649	0	---	91.2622
nonsynony	Olfir494:NM_146737:exon1:c.G295T:p.A99S,	chr7	1.16E+08	1.16E+08	G	T	0,0,68,134	0,0,473,0	0.663366	0.507463	0	---	89.86193
nonsynony	Leng1:NM_027203:exon2:c.A281G:p.K94R,	chr7	3615315	3615315	T	C	0,122,0,0	0,0,0,190	1	0.581967	0	---	89.34221
synonymou	Myo1b:NM_001161817:exon2:c.G75A:p.P25P,Myo1b:NM_010863:exon2	chr1	51939987	51939987	C	T	0,66,0,102	0,283,0,1	0.607143	0.264706	0.003521	0	53.86659
nonsynony	Mast3:NM_199308:exon2:c.G121A:p.V41M,	chr8	73313953	73313953	C	T	0,0,0,93	0,270,0,0	1	0.344086	0	---	88.38949
nonsynony	Arhgef40:NM_001145921:exon5:c.T1723G:p.S575A,Arhgef40:NM_1982	chr14	52610695	52610695	T	G	0,0,101,0	0,0,0,217	1	0.514851	0	---	85.00492
nonsynony	Ajap1:NM_001099299:exon1:c.C13A:p.Q5K,	chr4	1.53E+08	1.53E+08	G	T	0,0,60,137	0,0,382,0	0.695431	0.408759	0	---	84.79434
nonsynony	Mug1:NM_008645:exon2:c.G97C:p.V33L,	chr6	1.22E+08	1.22E+08	G	C	0,98,0,0	0,0,218,0	1	0.673469	0	---	83.66294
nonsynony	Fam71e2:NM_172895:exon9:c.A1859G:p.E620G,	chr7	4709455	4709455	T	C	0,123,0,0	0,0,0,163	1	0.382114	0	---	83.55343
synonymou	Aox3:NM_023617:exon7:c.G525A:p.Q175Q,	chr1	58184488	58184488	G	A	28,0,58,0	0,0,159,1	0.325581	0.464286	0	---	14.18689
nonsynony	Dmbt1:NM_007769:exon36:c.G3836A:p.R1279H,	chr7	1.38E+08	1.38E+08	G	A	144,0,61,0	0,0,335,0	0.702439	0.298611	0	---	81.60544
nonsynony	Olfir750:NM_207558:exon1:c.G887T:p.R296L,	chr14	51690180	51690180	C	A	105,0,0,0	0,187,0,0	1	0.238095	0	---	81.51921
nonsynony	Rab3gap2:NM_001163754:exon26:c.C3038T:p.A1013V,	chr1	1.87E+08	1.87E+08	C	T	0,0,0,114	0,173,0,1	1	0.236842	0.005747	0	80.5831
nonsynony	Olfir345:NM_146945:exon1:c.A565T:p.S189C,	chr2	36496125	36496125	A	T	54,0,0,124	366,0,0,0	0.696629	0.395161	0	---	79.17321
nonsynony	Sbsn:NM_172205:exon1:c.G616T:p.G206C,	chr7	31537196	31537196	G	T	0,0,76,143	0,0,360,0	0.652968	0.566434	0	---	78.99142
nonsynony	1600015110Rik:NM_001081273:exon1:c.G431T:p.G144V,	chr6	48880497	48880497	G	T	0,0,16,128	0,0,213,0	0.888889	0.734375	0	---	78.9832
nonsynony	Vmn2r15:NM_001104626:exon3:c.C1228G:p.H410D,	chr5	1.1E+08	1.1E+08	G	C	0,158,106,	0,0,374,0	0.598485	0.449367	0	---	77.72261
nonsynony	Dscaml1:NM_001081270:exon29:c.G5110A:p.V1704M,	chr9	45557744	45557744	G	A	80,0,27,0	0,0,448,0	0.747664	0.425	0	---	72.89168
nonsynony	Syn2:NM_013681:exon11:c.C1321A:p.L441I,Syn2:NM_001111015:exon	chr6	1.15E+08	1.15E+08	C	A	87,0,0,0	0,185,0,0	1	0.643678	0	---	72.75217
nonsynony	Fbxw22:NM_001014395:exon4:c.G301T:p.A101S,	chr9	1.09E+08	1.09E+08	C	A	157,97,0,0	0,309,0,0	0.61811	0.343949	0	---	71.22592
nonsynony	Gpr110:NM_133776:exon12:c.G1786T:p.A596S,	chr17	43447608	43447608	G	T	0,0,73,165	0,0,246,0	0.693277	0.545455	0	---	70.98739
synonymou	Abca12:NM_175210:exon29:c.C4365A:p.L1455L,	chr1	71335129	71335129	G	T	0,0,9,22	0,0,54,0	0.709677	0.636364	0	---	12.79731
nonsynony	Got1l1:NM_029674:exon8:c.C1036G:p.R346G,	chr8	28308436	28308436	G	C	0,116,0,0	0,0,124,0	1	0.741379	0	---	70.90091



全遺伝子領域をシーケンスすると？

変異が他のがんに比べて多くはない血液がんですら...

人力



~ 2 週間

©イラストわんパグ



May 8, 2018



26,038報の論文で報告され
た遺伝子領域の 5,621,127
変異を人が論文を読んで
データベース化



これを医師が一つ一つ
確認している現実

- COSMIC: exploring the world's knowledge of somatic mutations in human cancer

<http://cancer.sanger.ac.uk/cosmic>



膨大な量の電子化知識の氾濫

人智・人力を超えた世界

- 米国NIHのPubMed(医学・生物系論文の要旨データベース)上には2018年までに2800万件の論文が登録。
印刷するとその厚さは、富士山の高さを超える(高度4km)。
- 2018年 がんに関する論文だけでも**20万報**を超えた。
- 論文数は指数関数的に増えており、2050年には大気圏外(**高度100km**)に達する高さになる。
- 2015年 薬の特許件数は1500万件以上。



Photo from Wikipedia

https://ja.wikipedia.org/wiki/富士山#/media/ファイル:MtFuji_FujiCity.jpg

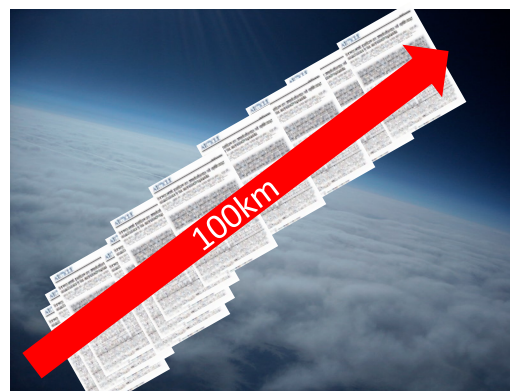
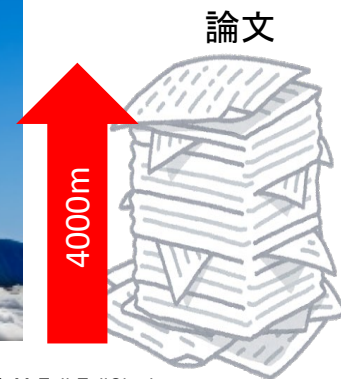


Photo by meteotek08, from flickr

<https://flic.kr/p/65stSJ>

CC BY-SA 2.0

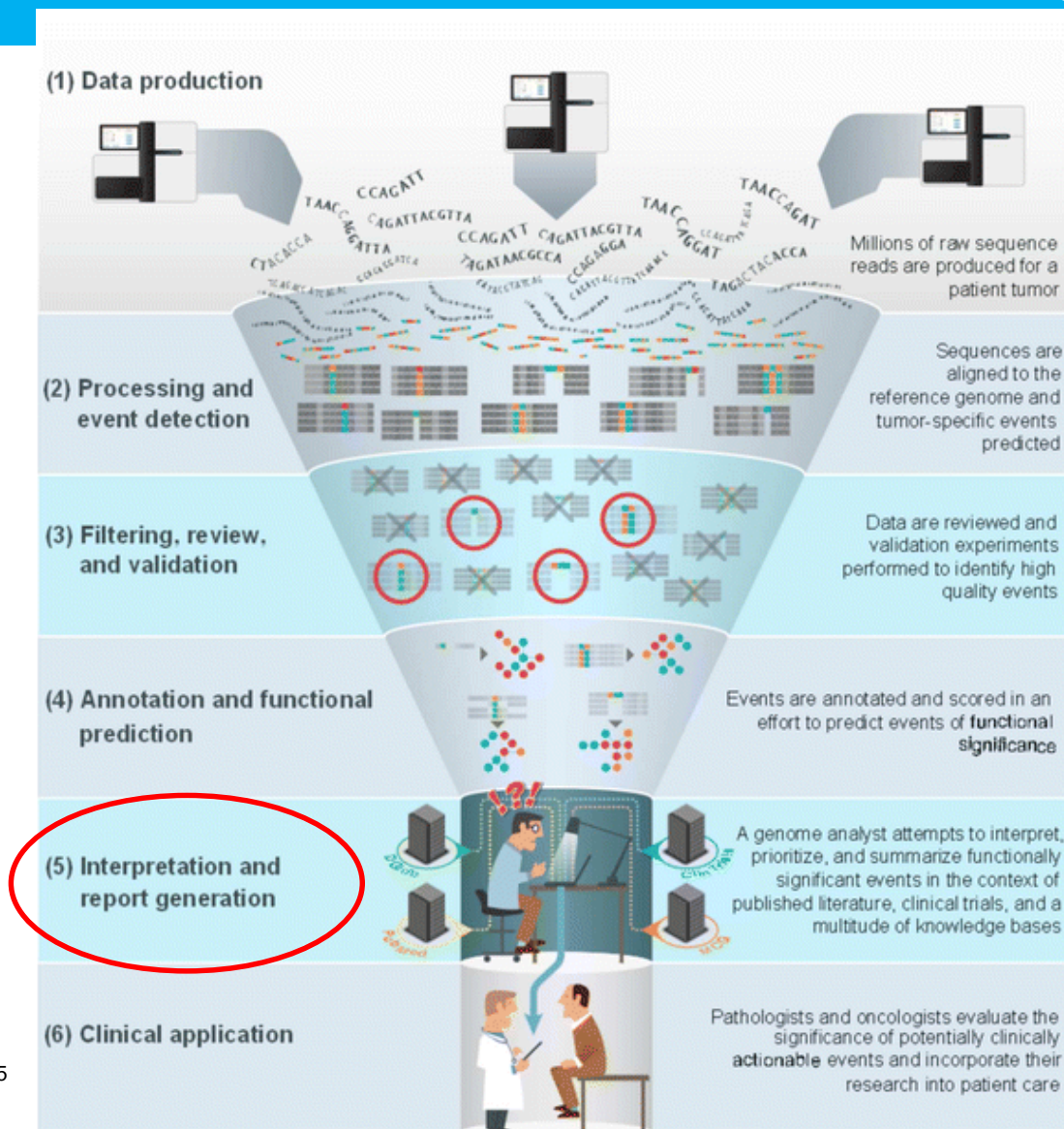
©いらすとや



臨床シーケンスのボトルネック



ゲノムの変異は
ぞろぞろ
見つかるが、
その解釈と翻訳が
ボトルネック



Good BM et al.
Organizing knowledge to enable personalization of medicine in cancer. *Genome Biol.* 2014 Aug 27;15(8):438. doi: 10.1186/s13059-014-0438-7. Fig.1
<https://genomebiology.biomedcentral.com/articles/10.1186/s13059-014-0438-7> CC BY 4.0

米国のトップがんセンターは人工知能を導入 Cognitive Computingが拓く未来

著作権等の都合により省略しました

IBM Watsonがクイズ番組に出場した際の画像
<https://www.ultius.com/ultius-blog/entry/could-artificial-intelligence-replace-human-writers.html>

米国のクイズ番組 Jeopardy! に出場
コンピュータの能力を立証

著作権等の都合により省略しました

CRAIN'S NEW YORK BUSINESS
IBM's Watson joins Genome Center to cure
cancer
2014年3月
http://www.crainsnewyork.com/article/20140319/HEALTH_CARE/140319845/ibms-watson-joins-genome-center-to-cure-cancer

New York Genome
Centerに脳腫瘍の医療の
サポートに導入された
Watson

- IBM Watson Oncologist
“Watson helping oncologists treat cancer patients”
- MD Anderson’s Oncology Expert Advisor
“Watson closing the loop of clinical trials, patient care, and treatment development”
- Mount Sinai
- Baylor College of Medicine
- Cleveland Clinic
- Mayo Clinic

他8カ所



医科研: ワトソン使用開始 ～がん研究は人知を越えた領域へ～



著作権等の都合により省略しました

日本経済新聞 2015年7月30日

がん最適治療 人工知能で 日本人向けに研究
日本IBM・東大

日本経済新聞
2015年7月30日朝刊

Watson Genomic Analytics
(ワトソン・ジノミック・アナリティクス)

現) Watson for Genomics

がん細胞のゲノム情報を解釈し
効果のあると期待できる抗がん剤
をその根拠(論文等)と共に
提示する。

Watson Early Adaptor Program
医科研は、北米以外で唯一の
採択機関

人工知能を臨床シーケンスに応用するための研究を行っています



医科研 全ゲノム臨床シーケンス



インターネットから切り離された パーソナルゲノム空間

Ion Proton
© Thermo Fisher Scientific

イルミナ株式会社 HiSeq 2500



© Quantum Biosystems Inc. 2019

3500 Genetic Analyzer
© Thermo Fisher Scientific

臨床シーケンスラボラトリー



生体認証システム

高セキュリティ パーソナルゲノム 専用スパコン領域

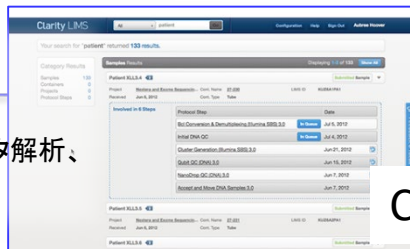
© ヒトゲノム解析センター



ソフトウェアのバージョン、パラメータ、
実行者など全ての入出力に関するログを自動的に記録

Laboratory Information Management System (LIMS)

- サンプル調整からデータ解析、
解析結果まで一元管理
- トレーサビリティを担保



株式会社ワールドフュージョン
BaseSpace Clarity LIMS

Clarity LIMS

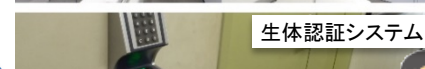
データ解析室



USBなど出力端子無し



全室モニターシステム



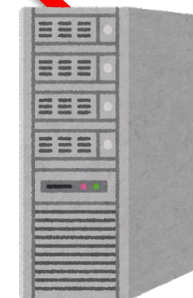
生体認証システム

臨床系研究室へは
VPN・シンクライアント接続

© イラストわんぱく

コントロール室

(別系統で管理)



© いらすとや



IBM “Watson for Genomics” を用いた 臨床シーケンス研究



がん細胞で起こっている変異
1,000～100,000



日本アイ・ビー・エム株式会社
IBM Watson

アライメント
変異解析
on Shirokane3



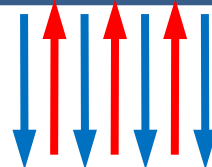
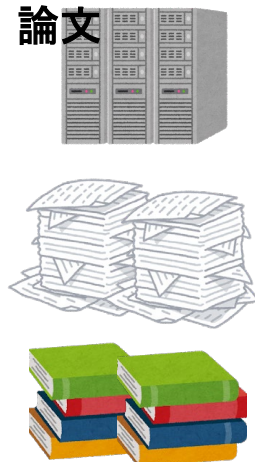
©ヒトゲノム解析センター



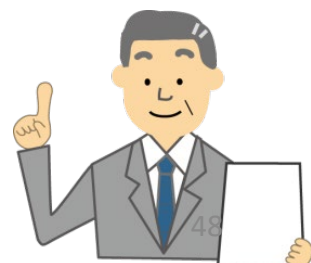
©いらすとや

SOFTLAYER

データベース
出版された
論文



効果が期待できる
抗がん剤のサマリー



©イラストわんパグ

Watson for Genomics



全遺伝子領域をシーケンスすると？

～スピードと網羅性が臨床研究を変える～

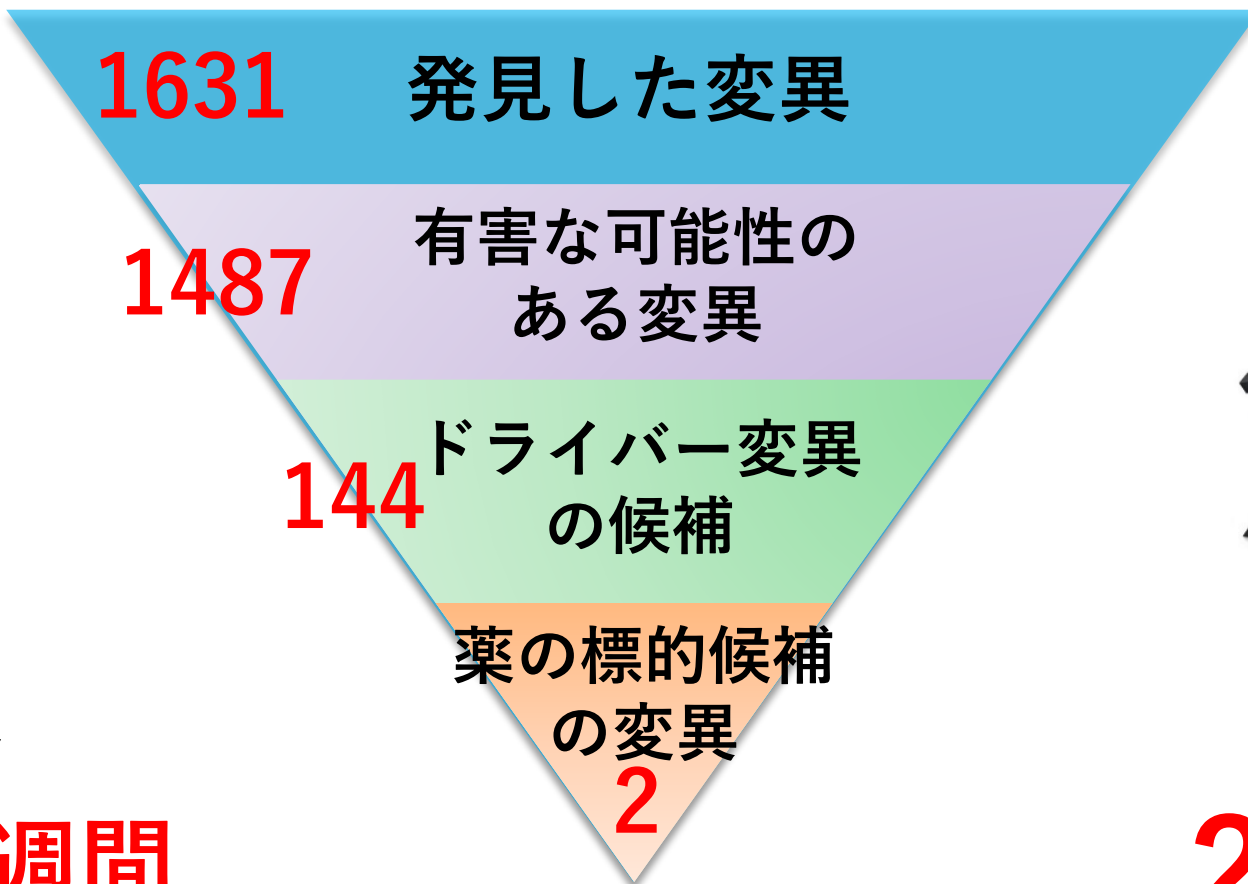


人力

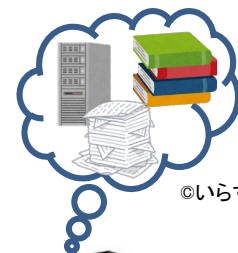


～ 2 週間

©イラストわんパグ



Watson



日本アイ・ビー・エム株式会社
IBM Watson

2分



大腸がん細胞株を使ったWatsonの検証

既に変異がわかっており原因が知られている大腸がん細胞株データを使って



• 大腸がん細胞株: RKO

- List of genomic variations is publicly available
 - Mouradov et al., *Cancer Research* (2014)
 - 一塩基変異 (SNVs)



Reference: ...ACAT**C**GGCA...
Observed: ...ACAT**A**GGCA...

• Watson for Genomics へのインプットファイル

- 4,237 SNVs のリスト
- VCF format





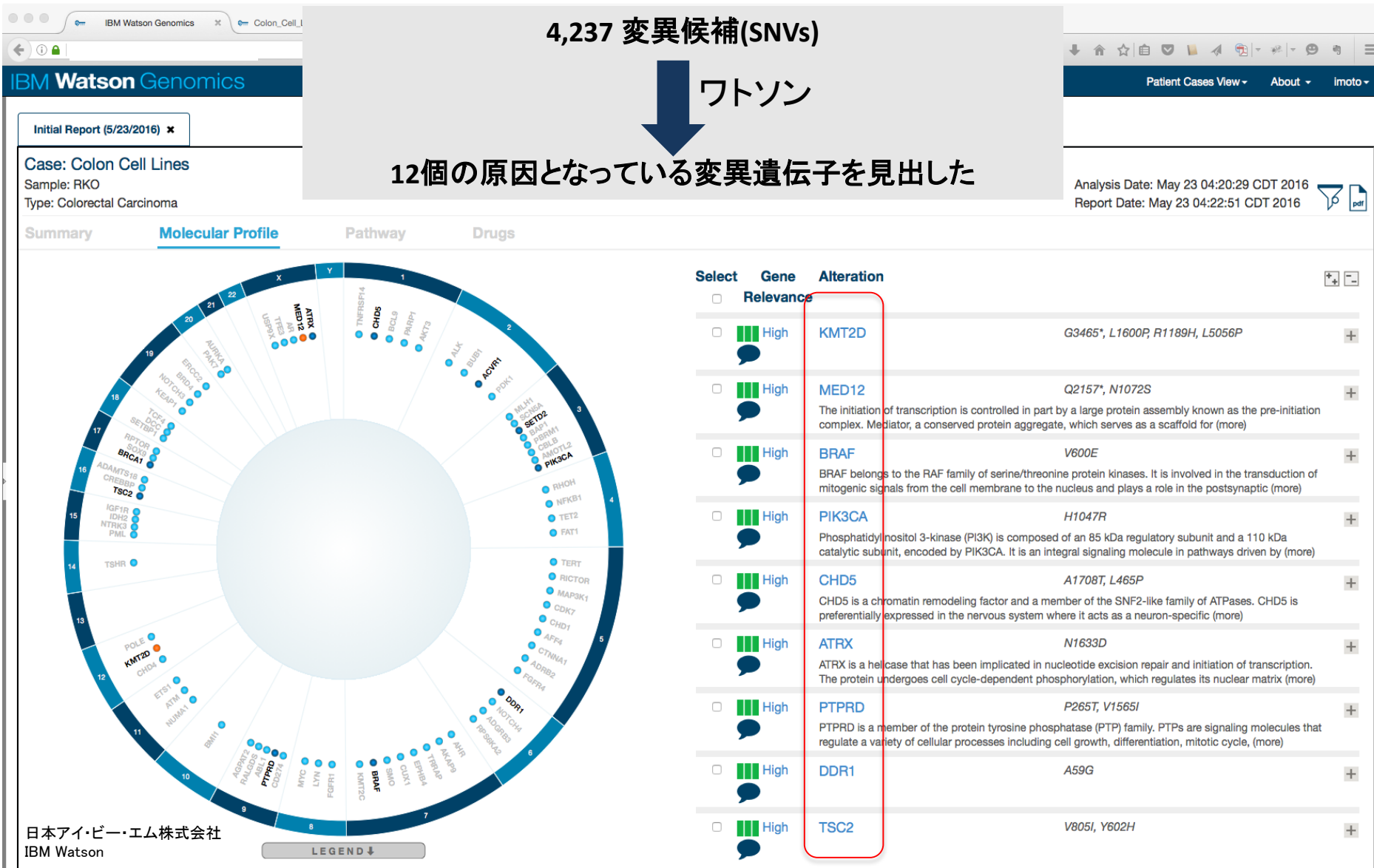
がんの原因: 変異遺伝子を探索



4,237 変異候補(SNVs)

ワトソン

12個の原因となっている変異遺伝子を見出した



IBM Watson Genomics

Patient Cases View About imoto

Initial Report (5/23/2016) x

Watson が提示す 3つの薬剤カテゴリー

Case: Colon Cell Lines
Sample: RKO
Type: Colorectal Carcinoma

Analysis Date: May 23 04:20:29 CDT 2016
Report Date: May 23 04:22:51 CDT 2016

薬とその標的遺伝子を提示

Summary Molecular Profile Pathway Drugs

Relevance	Alteration	Approved for Colorectal Carcinoma	Investigational for Colorectal Carcinoma	Off Label
■■■	BRAF V600E	Regorafenib	MLN2480 Vemurafenib Cobimetinib(GDC-0973 XL518) PD0325901 Binimetinib(MEK162) Trametinib	Dabrafenib
■■■	PIK3CA H1047R	-	Buparlisib(BKM120) Alpelisib(BYL719) MLN1117	-
■■■	BRCA1 D435Y	-	Talazoparib(BMN 673) Olaparib(AZD-2281)	-
■■■	ATM D2913G	-	Talazoparib(BMN 673) Olaparib(AZD-2281) AZD6738	-
■■■	RPTOR A245T	-	AZD2014 INK128 Sirolimus Everolimus Temsirolimus	-
■■■	IGF1R M122I	-	Ganitumab(AMG 479)	-
■■■	DDR1 A59G	-	Imatinib Dasatinib	Nilotinib
■■■	AURKA L222P	-	Alisertib(MLN8237)	-

変異の入っている標的遺伝子(左)と変異体の名称(右)

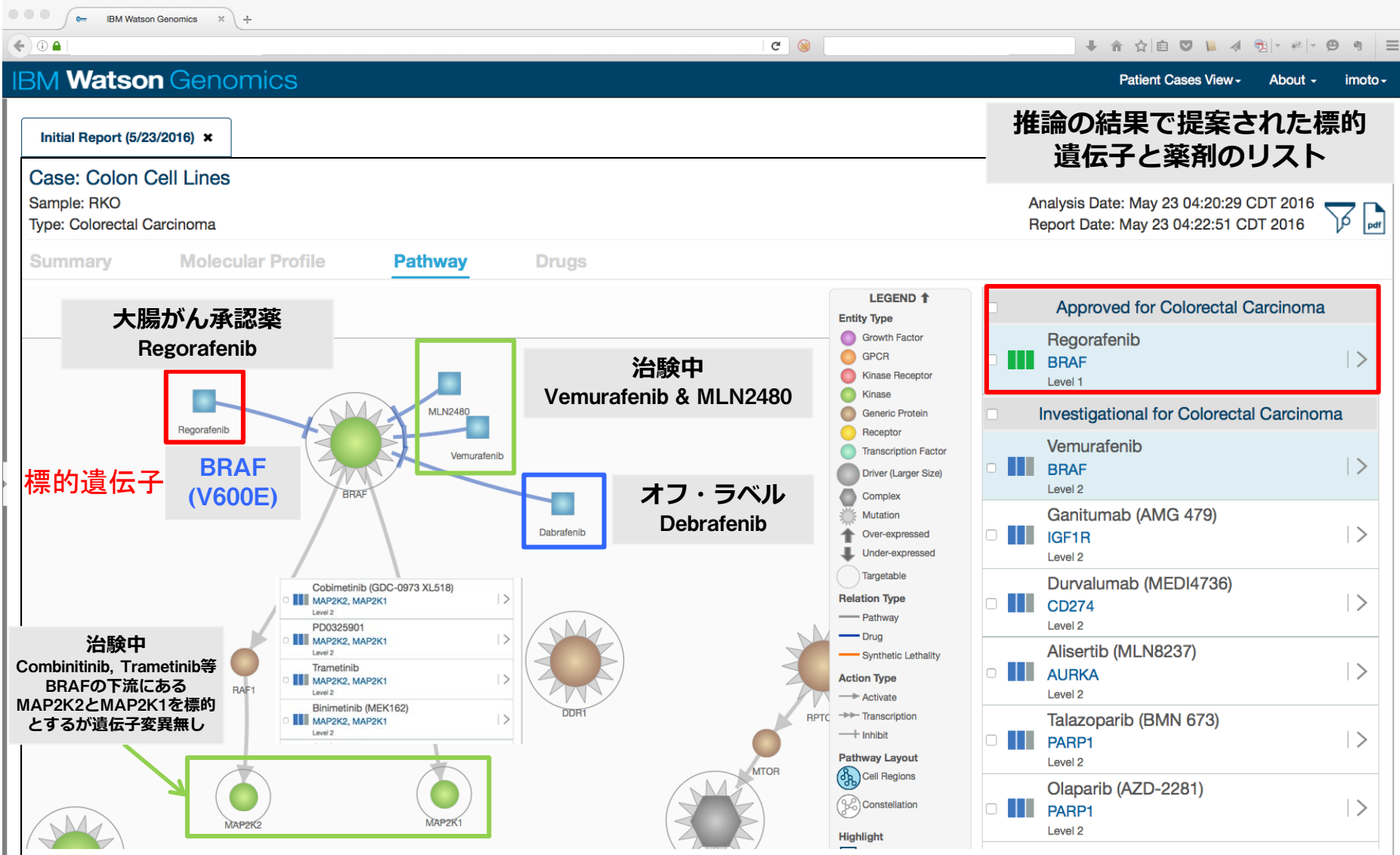
大腸がんに対して米国で承認されている薬剤
(注：日本では未承認)

大腸がんに対して米国で現在臨床試験中の薬剤

オフ・ラベル：他のがんについて米国で承認されている薬剤



薬とその標的遺伝子を提示した根拠 その論文・薬剤情報へのリンク





日本の臨床試験



国立保健医療科学院

臨床試験ポータルサイト

<https://rctportal.niph.go.jp/s/>

UMIN 臨床試験登録システム
(UMIN: 国立大学附属病院長会議)

29,916件 (2018/2/1時点)

JapicCTI (JAPIC: 財団法人日本医薬情報センター)

3,870件 (2018/5/19時点)

治験促進センター 臨床試験登録システム(JMA: 日本医師会)

324件 (2018/2/1時点)

ClinicalTrials.gov (米国臨床試験ポータルサイト)

XML全データ18万件



臨床研究情報ポータルサイト 医療関係者の方へ
<https://rctportal.niph.go.jp/s/>

webでは全19,483件と出るが、
英↔日、日↔日の重複を除いた



日本へのローカライゼーション



Summary Drugs Molecular Profile

2B Olaparib

2B Rucaparib Phosphate

2B Niraparib

3A Veliparib

3A Talazoparib

3A Alpelisib

3A Copanlisib

3A MLN1117

3A Everolimus

3A AZD4547

Over View Clinical Trials Clinical Trials (Watson) テキストでの絞り込み検索

Olaparib and BRCA and

Get Data

HELP

US Clinical Trials 7 studies

status filter	ID(URL)	title	phase	condition	min age	max age	gender	sponsor	募集状況	ID(URL)	試験名	対象疾患	フェーズ	年齢	性別	提供元
☐ Not yet recruiting ☒ Recruiting ☒ Enrolling ☒ Active, not recruiting ☒ Available ☐ Not recruiting	NCT01844986	Olaparib Maintenance Monotherapy in Patients With BRCA Mutated Ovarian Cancer Following First Line Platinum Based Chemotherapy.	Phase 3	Newly Diagnosed Advanced Ovarian Cancer FIGO Stage III-IV BRCA Mutation Complete Response Partial Response First Line Platinum Chemotherapy	18 Years		Female	AstraZeneca	☒ 募集開始前 ☒ 参加者募集中 ☒ 一般募集中 ☒ 限定募集中 ☐ 試験終了 ☐ 参加者募集終了 ☐ 試験継続中 ☐ その他	JapicCTI-132360	白金製剤による初回化学療法を受けたBRCA変異を有する卵巣癌患者にオラパリブを単独投与する	白金製剤による初回化学療法を受けたBRCA変異を有する卵巣癌患者	第Ⅲ相	18歳～	女	アストラゼネカ
Active		Olaparib Treatment in BRCA Mutated		Platinum Sensitive BRCA Mutated Relapsed Ovarian Cancer					参加者募		白金製剤による化学療法後に完全奏効又は部					

募集状況の絞り込み

元ページのリンク

Watson のアウトプットを
読み込み



日本アイ・ビー・エム株式会社
IBM Watson

臨床研究情報ポータルサイト 医療関係者の方へ
<https://rctportal.niph.go.jp/s/>

臨床研究情報ポータルサイト

患者様・ご家族など一般の方向け 医療関係者の方向け

検索

クリップした情報 (0件)

JAPIC試験ID JapicCTI-132360
最終情報更新日: 2016年9月29日 登録日: 2016年1月13日

白金製剤による初回化学療法を受けたBRCA変異を有する進行卵巣癌患者 (FIGO 進行期 分類3~4期) にオラパリブを維持療法として単独投与する無作為化二重盲検プラセボ対照多施設共同第3相試験

○ 基本情報

関連ID NCT01844986: ClinicalTrial.gov

○ 対象疾患

試験実施者 アストラゼネカ

試験の種類 介入試験(薬剤)

試験の概要 白金製剤による初回化学療法を受けたBRCA変異を有する卵巣癌患者にオラパリブを単独投与する

JAPICで詳細を確認する

この情報をクリップする

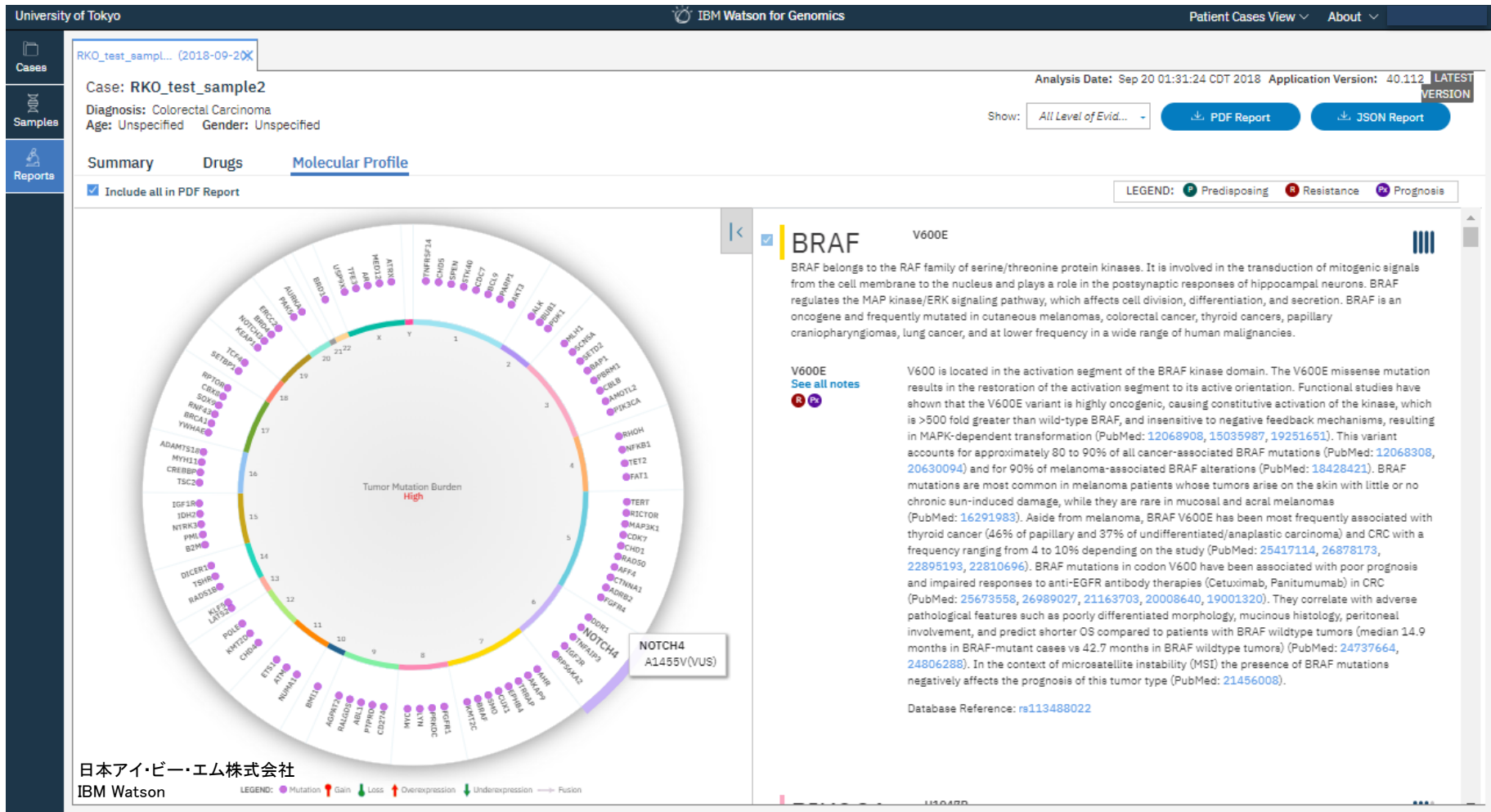
Update: 変異遺伝子を探索 (RKO)

Driver gene candidates narrowed down
from 4,237 mutations' list (SNVs)



2 highly relevant genes (BRAF, PIK3CA)
+ Tumor Mutation Burden (TMB)

2016年の解析では
12個の候補



Update: 効果が期待される薬剤

University of Tokyo IBM Watson for Genomics Patient Cases View About

Case: RKO_test_sample2
Diagnosis: Colorectal Carcinoma
Age: Unspecified Gender: Unspecified

Analysis Date: Sep 20 01:31:24 CDT 2018 Application Version: 40.112 LATEST VERSION

Show: All Level of Evid... PDF Report JSON Report

Summary Drugs Molecular Profile

High mutation burden, which is a strong predictive marker of response to treatment with immune checkpoint inhibitors, has been detected in this tumor. The V600E mutation in gene BRAF is a predictor of response to treatment with vemurafenib. In addition, mutations in BRAF resulting in activation of the MEK signaling pathway and mutations in RAS resulting in activation of the MAPK/ERK signaling pathway have been detected. Treatment options with drugs targeting one or both of these pathways may be therapeutic options for this patient.

NCCN-Compendium has listed V600E missense mutation and codon 600G to A mutation in BRAF as a predictor of poor prognosis in Colorectal Carcinoma.

Three categories of recommended drugs

Relevance	Alteration	Approved for: Colorectal Carcinoma	Investigational for: Colorectal Carcinoma	Approved for Other Indications
III	BRAF V600E	Approved for "Targeted" Cancer	Investigational: Under Clinical Trials	Off Label: Approved for "Other" Cancers
III	PIK3CA H1047R			
III	TMB High			

Regorafenib が無くなった



卵巢原発腹膜偽粘液腫症例

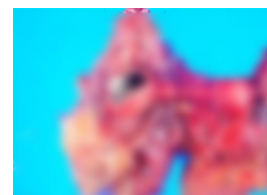


全ゲノムシーケンス解析と Watson for Genomic で 標的遺伝子と抗がん剤候補が見つかる

- Pseudomyxoma peritonei (**PMP**) : 稀な腫瘍で百万人に1~2人。腫瘍から産生される粘液で腹腔内にゼリー状物質が蓄積する。大部分が虫垂原発の腫瘍。緩序に進行するが、再発を起こしやすい。
- 【症例】 38歳女性 主訴は腹部膨満
CTにて卵巢腫瘍の腹膜偽粘液腫が疑われる。
完全減量手術+温熱化学療法(MMC)が行われた。
病理組織検査で卵巢成熟奇形腫による腹膜偽粘液腫と診断



古川洋一教授, MD, PhD
(東大医科研
ゲノム診療部)

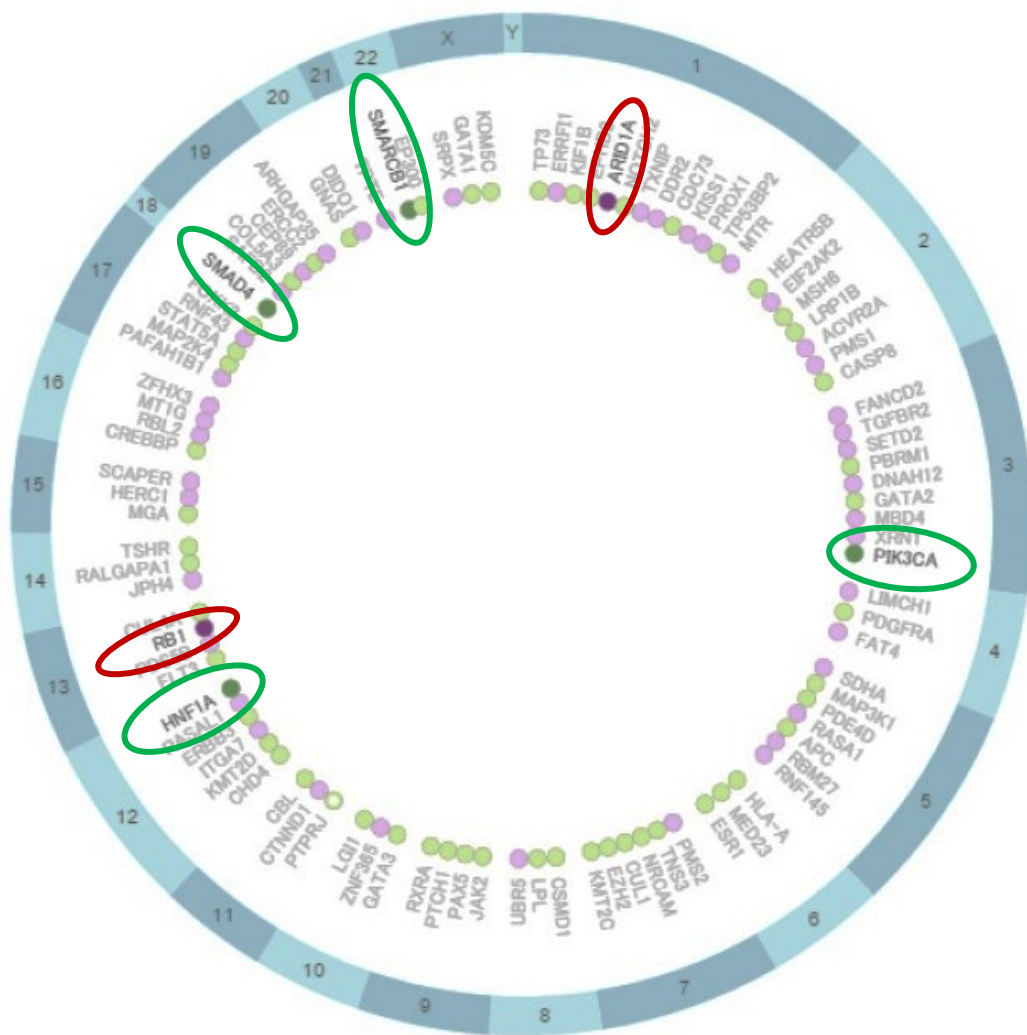


- 腫瘍組織を用いた全ゲノム解析を実施





解析結果 (Molecular Profile)

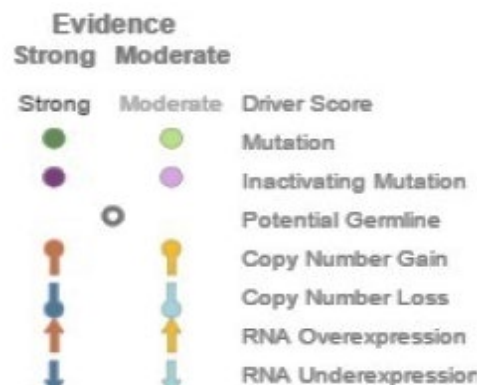


Watsonの推測:

- 250万個のゲノム変異から100遺伝子がドライバー遺伝子変異として同定された。
- これらの中で6遺伝子が strong evidenceとして予測された。



- Mutation: *SMAD4*, *SMARCB1*,
PIK3CA, *HNF1A*
- Inactivated mutation: *RB1*, *ARID1A*



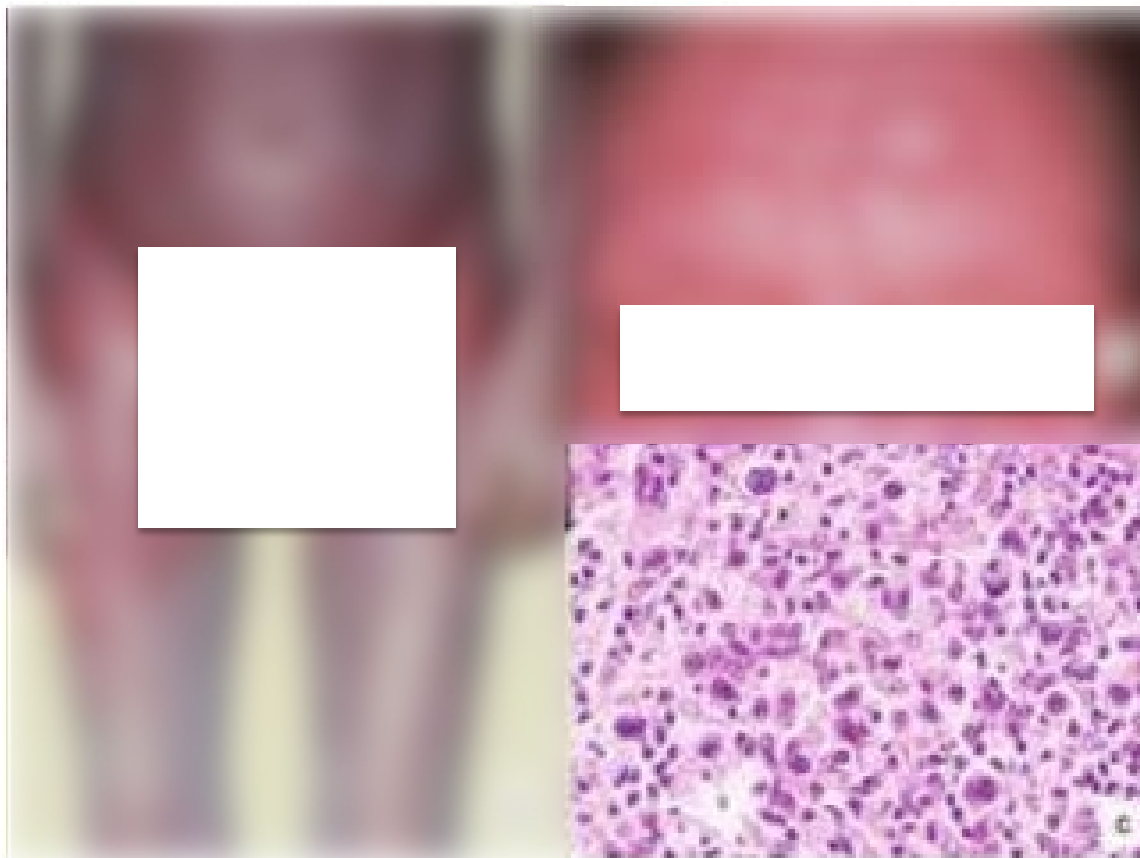


皮膚T細胞性リンパ腫



(Mycosis Fungoides and Sezary Syndrome)

70代男性





専門医の解釈



Whole Exome Sequence

専門医解釈 (全遺伝子シーケンス)

315 全遺伝子変異情報(患者皮膚サンプル)

314 ドライバー変異ではない変異

1 **ドライバー変異**

薬剤標的

0

ゲノムの異常からは
薬剤の候補は見つからなかった

Whole Transcriptome Sequence

専門医解釈 (全遺伝子発現解析)

約2万個 全遺伝子発現/パスウェイ情報
(患者皮膚サンプル)

薬剤標的ではない
異常

**薬剤標的と
なり得る
異常**

6個

RNA発現の異常からは
6つの薬剤の候補が見つかった



AI と専門医、解釈の比較



専門医の同定した6つの抗がん剤標的候補のうち4つは **Watson** も同定
更に、4つの標的候補を提示

**Database
Search
by MD**

6

8

Watson

2

4

4

CD30 (TNFRSF8)
CD25

CCR4
PDL1
CTLA4
LCK

BTK [1,3]
CDKN2B [1,2]
CDKN1C [1,2]
PIK3CD [1]

Watsonのcorpus
情報にない

↓
IBM ヘフィード
バックを行った

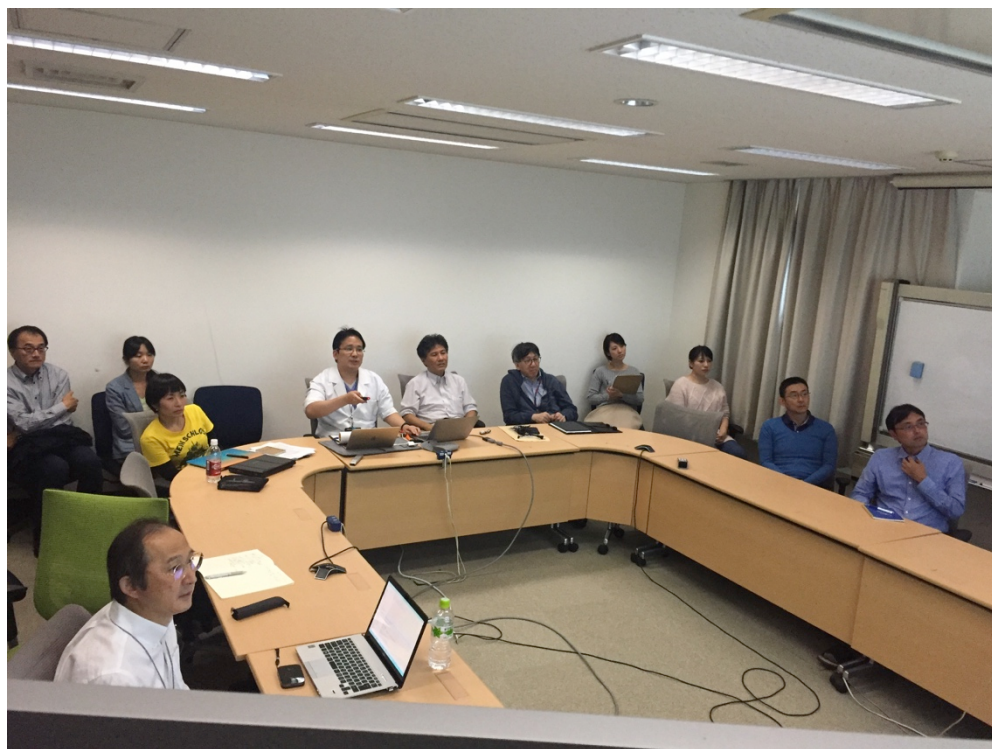
1. 文献根拠が乏しい
2. 米国でのみ使用可能
3. B細胞腫瘍にのみ発現



Watson の解釈を評価しフィードバック



1ヶ月に2回(隔週)で開催



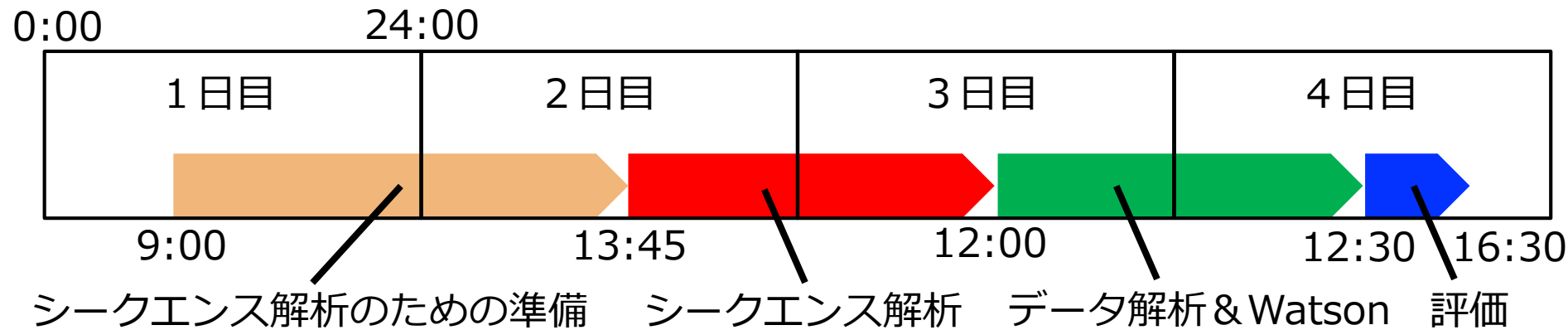
医師: 血液腫瘍内科、外科
臨床遺伝専門医
がん治療認定医
臨床薬理学
分子生物学
バイオインフォマティクス
スパコンシステム工学
生命倫理学
等のエキスパート(34名)



全ゲノム臨床シーケンス

40代 男性 急性骨髄性白血病

2017年12月



Turn Around Time = 3日と7時間半

7,596個のがん特異的ゲノム異常 (内 108 個の構造異常)
Long read nanopore

アライメント : CPUで数時間

➡ GPU (v100x8) で34分

**Shirokane5に
80基搭載**

医科研における AI を用いた臨床シーケンスが NHK ニュース（ニュース7）にて取り上げられました

COI STREAM にて医科学研究所では、人工知能を用いたゲノム情報の解釈法の開発をテーマの一つに掲げ、日本 IBM と研究開発を進めています。

著作権等の都合により省略しました

2016年8月4日付 NHK全国ニュース

人工知能 “特殊な白血病” 見抜き女性患者の命救う

64歳の女性が急性骨髄性白血病を発症。

標準的な化学療法後に致命的な造血不全に陥った。原因がわからず造血不全による肺炎なども併発し、次の治療方針も定まらなかった。そこで白血病細胞のDNAを次世代シーケンサーとスーパーコンピュータ、Watsonを用いて解析した結果、2次性白血病の可能性が高いと判断。

治療方針を変更した処、造血不全が回復。白血病も克服。

攻殻機動隊 Realized Project 受賞

2017年3月26日

審査員特別賞は「人工知能が、特殊な白血病患者の病名を10分ほどで見抜き、その生命を救った」で、**東京大学医科学研究所臨床シーケンス研究チームが受賞**した。

「Watson」は膨大な数の研究の論文の要約、薬の特許情報、さらにこれまでの医師が行った対応など、様々なデータを抽出し医師をバックアップ、患者の症状や対処法を導き出す手助けをする。そしてその結果も学習し、さらにデータベースとして確固たるものに成長するという。遺伝子解析と、対処法の提示としての“量とスピード”においては、人間を遙かに超える力を持ったシステムとなっているという。

(Impress 記事より)

著作権等の都合により省略しました

Gamewatch 「AnimeJapan 2017」
「攻殻機動隊 REALIZE PROJECT」最新活動を発表
2017年3月26日

<https://game.watch.impress.co.jp/docs/news/1051368.html>

厚生労働省通知

医政医発1219第1号

人工知能（AI）を用いた診断、治療等の支援を行うプログラムの利用と医師法第17条の規定との関係について

近年、機械学習の技術の進歩等により、診療を行うに当たって人工知能(AI)を用いた診断・治療支援を行うプログラムが用いられる機会が増加しており、今後、その果たす役割はますます大きくなるものと予想されている。

このような中、平成 29 年度厚生労働行政推進調査事業費補助金により、「AI 等の ICT を用いた診療支援に関する研究」が行われ、本研究の報告書が取りまとめられたところである(概要は別添参照)。

人工知能(AI)を用いた診断・治療支援を行うプログラムを利用して診療を行う場合についても、診断、治療等を行う主体は医師であり、医師はその最終的な判断の責任を負うこととなり、当該診療は医師法(昭和 23 年法律第 201 号)第 17 条の医業として行われるものであるので、十分ご留意をいただきたい。

研究班（全員が東大医科研にてゲノムとAIに取り組む研究者）

横山 和明 東京大学・医科研附属病院 血液腫瘍内科 助教

井元 清哉 東京大学・医科研 健康医療データサイエンス分野 教授

古川 洋一 東京大学・医科研 臨床ゲノム腫瘍学分野 教授

湯地 晃一郎 東京大学・医科研 国際先端医療社会連携研究部門 特任准教授



医政医発 1219 第 1 号
平成 30 年 12 月 19 日

各都道府県衛生主管部(局)長 殿

厚生労働省医政局医事課長
(公 印 省 略)

人工知能 (AI) を用いた診断、治療等の支援を行うプログラムの利用と
医師法第 17 条の規定との関係について

近年、機械学習の技術の進歩等により、診療を行うに当たって人工知能 (AI) を用いた診断・治療支援を行うプログラムが用いられる機会が増加しており、今後、その果たす役割はますます大きくなるものと予想されている。

このような中、平成 29 年度厚生労働行政推進調査事業費補助金により、「AI 等の ICT を用いた診療支援に関する研究」(研究代表者：横山和明東京大学医科学研究所附属病院血液腫瘍内科助教)が行われ、本研究の報告書が取りまとめられたところである(概要は別添参照)。

当該報告書では、人工知能 (AI) を用いた診断、治療等の支援を行うプログラムを利用して診療を行うことについて、本研究において行った AI 等の ICT を用いた診療支援に関する調査等を踏まえ、「AI は診療プロセスの中で医師主体判断のサブステップにおいて、その効率を上げて情報を提示する支援ツールに過ぎない」、「判断の主体は少なくとも当面は医師である」等と整理している。

上記のとおり、人工知能 (AI) を用いた診断・治療支援を行うプログラムを利用して診療を行う場合についても、診断、治療等を行う主体は医師であり、医師はその最終的な判断の責任を負うこととなり、当該診療は医師法(昭和 23 年法律第 201 号)第 17 条の医業として行われるものであるため、十分ご留意をいただきたい。

貴職におかれては、内容を御知知いただくとともに、貴管下保健所設置市、特別区、関係機関及び関係団体等に周知をお願いする。

(別添)

厚生労働行政推進調査事業費補助金(地域医療基盤開発推進研究事業)
AI 等の ICT を用いた診療支援に関する研究(II29-医療-指定-015)

<研究代表者>

横山 和明 東京大学・医科学研究所附属病院 血液腫瘍内科 助教

<研究分担者>

井元 清哉 東京大学・医科学研究所 健康医療データサイエンス分野 教授
古川 洋一 東京大学・医科学研究所 臨床ゲノム腫瘍学分野 教授
湯地 晃一郎

東京大学・医科学研究所 国際先端医療社会連携研究部門 特任准教授

<研究要旨>

人工知能(AI)の利活用は医療分野に大きな影響を及ぼすと予想される。医師による診療プロセスの一部に AI が介在することは、医療の質の向上に貢献し得る一方で、新たな社会的・法的問題を生じる潜在性がある。また、診療のどのプロセスに AI が介在しているかは必ずしも明らかではない。さらに国内における AI 等の ICT を用いた診療支援等の研究状況も明らかでなく、これらの点について政策立案のための論点整理を行うことが必要である。本研究では、国内での AI を用いた診療支援研究について、有識者へのヒアリングを交えながら、診療のプロセスという観点からそれらの類型化を試み、以下の結果を得た。1)AI は診療プロセスの中で医師主体判断のサブステップにおいて、その効率を上げて情報を提示する支援ツールに過ぎない。2)AI には知識量の制約がなく、医師主体判断のサブステップにおいて、医師にデバイスによる気づきを与え得る。AI と医師との協働は医療の質向上に有用であると考えられる。3)AI の推測結果には誤りがあり得るが、判断の主体である医師が AI を用いた診療の責任を負うべきである。その前提として医師に対して AI についての適切な教育を行うべきである。4)本邦における AI による診療支援研究はまだ萌芽期段階である事、判断の主体は少なくとも当面は医師である事実を鑑みると、その規制の議論は時期尚早である。寧ろ保険医療分野における AI 開発に関わる医師および研究開発者などの人材育成と公的な支援体制の整備の方が優先されるべきである。

人工知能(AI)を用いた診断、治療等の支援を行うプログラムの利用と医師法第 17 条の規定との関係につ
厚生労働省 医政医発1219第1号 平成30年12月19日
<https://www.mhlw.go.jp/content/10601000/000468150.pdf>



医師と AI の協調 画像診断の誤診率が85%減



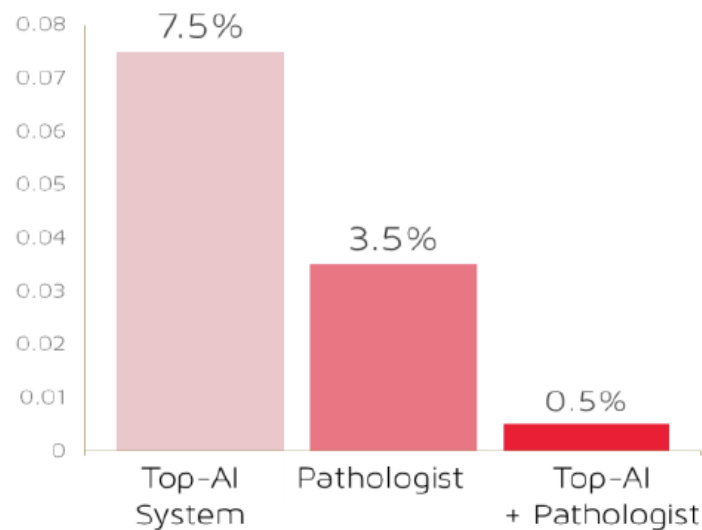
ハーバード大
& MIT

- 2016年実施の転移性乳癌 Camelyon Grand Challenge AI の誤診率は7.5% (1-AUC)
- 同じ標本セットを診断した人間病理医の誤診率は3.5%
- AIと医師診断の併用 ➡ 誤診率は0.5% に減少(85%減)

著作権等の都合により省略しました

転移性乳癌の画像

Dayong Wang et.al,
Deep Learning for Identifying Metastatic Breast
Cancer, 2016 Figure 2
<https://arxiv.org/pdf/1606.05718v1.pdf>



AI significantly reduces pathologist error rate in the identification of metastatic breast cancer from sentinel lymph node biopsies.



がんゲノム医療@日本



がんゲノム医療中核拠点病院(11カ所)



がんゲノム医療連携病院(100ヶ所)

厚生労働省
がんゲノム医療推進に向けた取組
<https://www.mhlw.go.jp/content/10901000/000341604.pdf> p13

著作権等の都合により省略しました、
がんゲノム医療、検査体制を拡充へ
厚生労働省

2018年11月29日 共同通信
<https://www.47news.jp/3027755.html>

がんゲノム医療

一般的に
対象となる人や状態

標準治療終了後

標準治療がない
原発不明がん、
希少がんなど

がん遺伝子パネル検査

多数の遺伝子を
同時に検査

検査結果

治療方針決定

a遺伝子

b遺伝子

c遺伝子

d遺伝子

...

...

d遺伝子変異あり

〈期待できる薬がある場合〉
d遺伝子変異に効果が期待できる薬の使用を検討(臨床試験など)

〈期待できる薬がない場合〉
その他の治療を検討

遺伝子変異なし

その他の治療を検討

国立がん研究センター がん情報サービス
がんゲノム医療 もっと詳しく知りたい方へ 図2 がん遺伝子パネル検査(がんゲノム医療)
https://ganjoho.jp/public/dia_tre/treatment/genomic_medicine/genmed02.html

← 医療現場でゲノム情報の臨床翻訳(ゲノムの変異情報を臨床的アクションに繋げる)ができるバイオインフォマティシャンは日本に40人もいません。世界的な人材不足。



Take Home Message



- 医科研では、2015年7月より、WfG をがん研究に活用してきました。これまでに
人工知能の情報の網羅性、アウトプットまでのスピードを実体験し、ゲノム医療における「未来」を感じている。
- 一般に認知されていることとの差異
 - WfG は診断はしない(医師を介さずに診断する AI (FDA Approved) も出てきている)。
 - 解釈結果はアップデートされる。
 - 誤りもある。
- データ解析の能力を有する人材、AIを使いこなす人材が、医科学・医療に不可欠になってきた。**人材育成が鍵**
- 「がん」での事例を他疾患に広げる時期が来ている