

本邦の遺伝統計学と 今後の国際戦略

岡 田 随 象

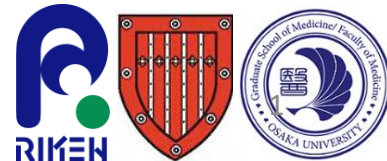
東京大学大学院医学系研究科 遺伝情報学

大阪大学大学院医学系研究科 遺伝統計学

理化学研究所生命医科学研究センター システム遺伝学チーム

URL : <http://www.sg.med.osaka-u.ac.jp/index.html>

E-mail : yuki-okada@m.u-tokyo.ac.jp

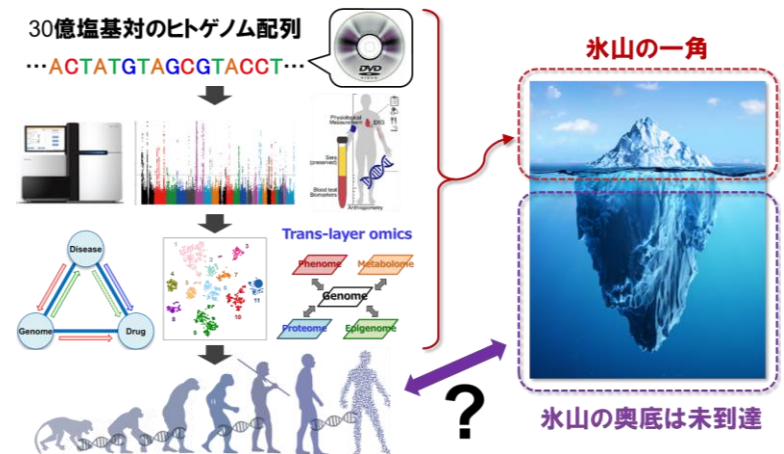
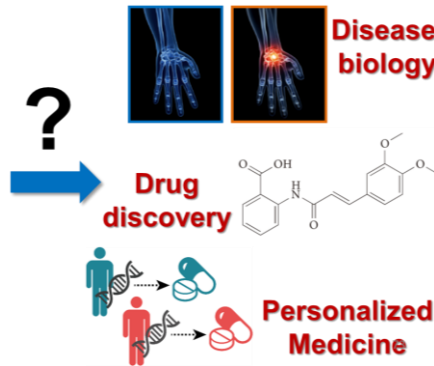
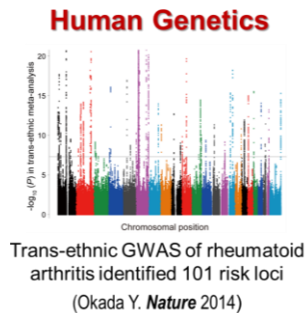


遺伝情報と形質情報の関わりを検討する遺伝統計学を主軸に、多彩なヒトオミクス情報のデータ解析を通じて、**疾患病態解明・ゲノム創薬・個別化医療**の実現に取り組んでいます。

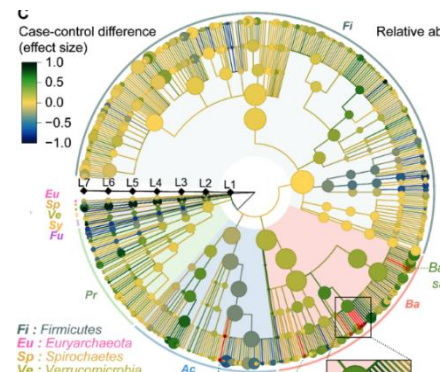
e-mail: yuki-okada@m.u-tokyo.ac.jp web: <http://www.sg.med.osaka-u.ac.jp/index.html> 場所: 医学部・研究教育棟6F北

ゲノム解析の成果をどのように活用すれば
病態解明・創薬・個別化医療に貢献できるのか
ほとんどわかっていない

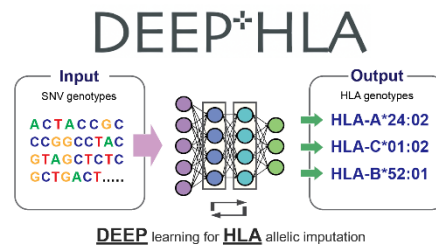
氷山の一角: ヒトゲノム配列のほとんどは未解明



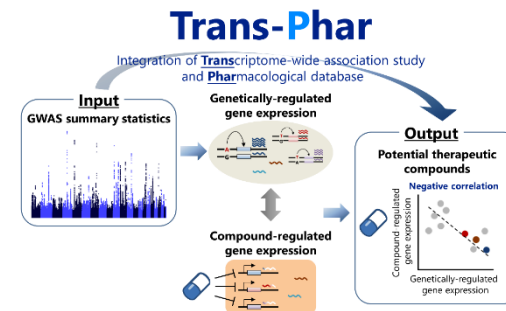
腸内微生物叢メタゲノム
ショットガンシーケンス



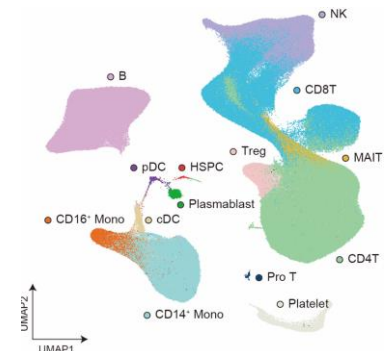
集団ゲノム情報への
畳み込み深層学習
(DEEP*HLA)



疾患ゲノム創薬の実装
(Trans-Phar)



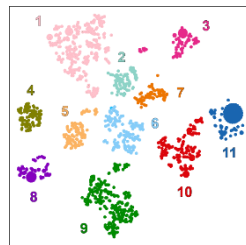
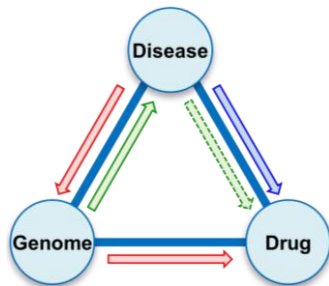
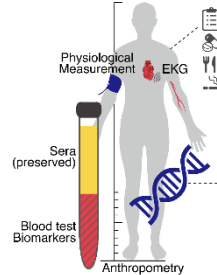
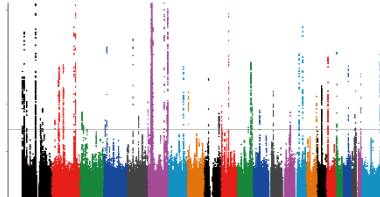
シングルセル解析
による疾患病態解明



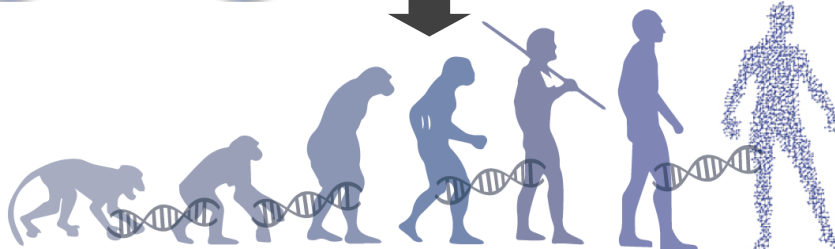
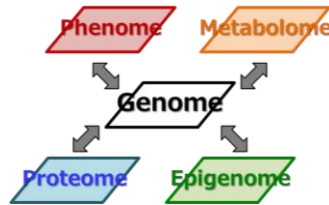
氷山の一角: ヒトゲノム配列のほとんどは未解明

30億塩基対のヒトゲノム配列

...ACTATGTAGCGTACCT...



Trans-layer omics



?

氷山の一角

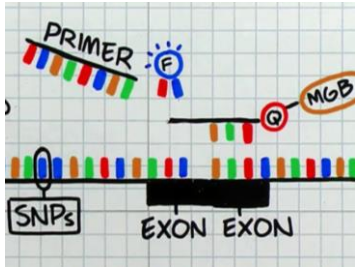


氷山の奥底は未到達

- 30億塩基対のヒトゲノム配列はヒトの設計図。そのメカニズムを解明すれば生命現象を再構成できるはずだが、ほとんどは**未解明**。
- 研究テーマ: 如何にヒトゲノム配列に隠された生命現象を引き出すか。

Wet・Dry双方の最新解析技術の先進的導入

TaqManアッセイ



SNPマイクロアレイ



ショートリードNGS



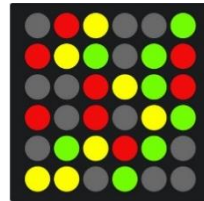
ロングリードNGS



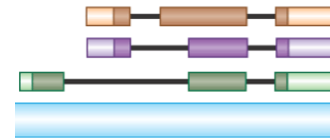
RT-PCR



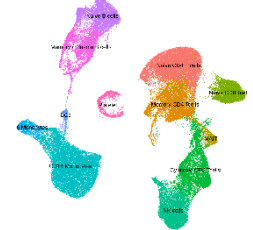
遺伝子マイクロアレイ



RNA-seq



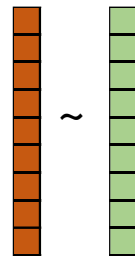
シングルセル解析



分割表検定

	AA	AG	GG
Case	360	480	160
Control	250	500	250

回帰分析

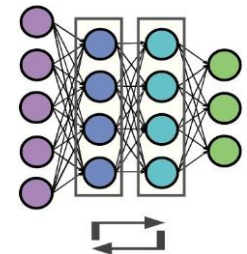


機械学習

Variants (90 K)	
ACTGAACGCAAAC	T
CCAGTATTCTACCT	T
CATGACTGCAAAC	T
ACAGTATGCTACAT	T
CCTAAATGCTACCT	T
AAAGTCTTCTACCT	T
ACTATATGCTACCT	T

Individuals (170 K)

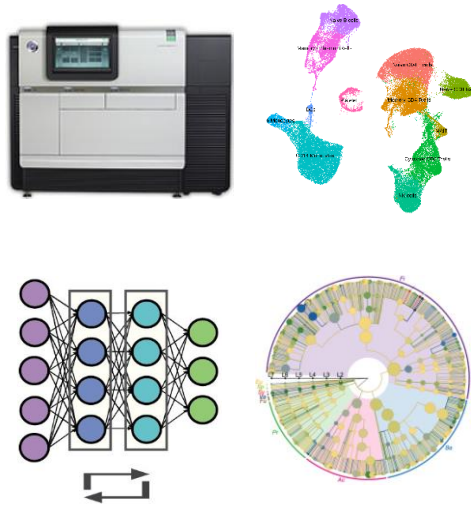
深層学習



- ・観測・解析技術の進歩は、常に予想を上回る速度で進む。
- ・Wet・Dry双方の最新解析技術の先進的導入が生命科学に不可欠。
- ・実験・解析原理の正確な理解が、革新的な研究を可能にする。

Curiosity, Methodology, Missionの融合イニシアティブ

Methodology-driven イニシアティブ



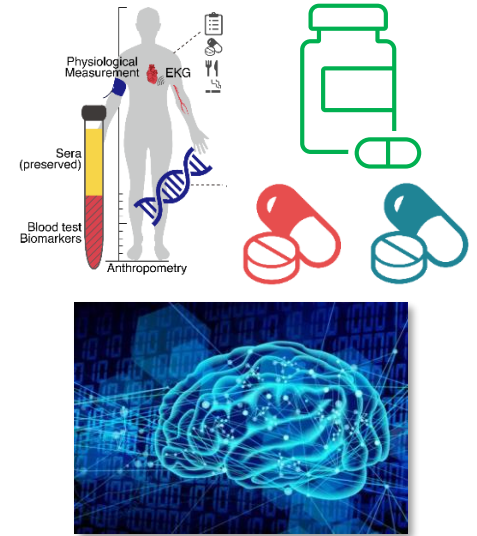
*We have more Data
than we can analyze.*

Curiosity-driven イニシアティブ



*Why want to climb Everest?
- Because it is there.*

Mission-oriented イニシアティブ



Hippocratic Oath

- **Curiosity**: 基礎研究の根幹を成す、生命現象解明への探求心。
- **Methodology**: 最新解析技術が可能にする、生命科学の新展開。
- **Mission**: 医学研究として不可欠な、疾病克服・健康長寿への挑戦。
- 3者の融合を通じて、**ライフサイエンスのイニシアティブ**を探求していく。

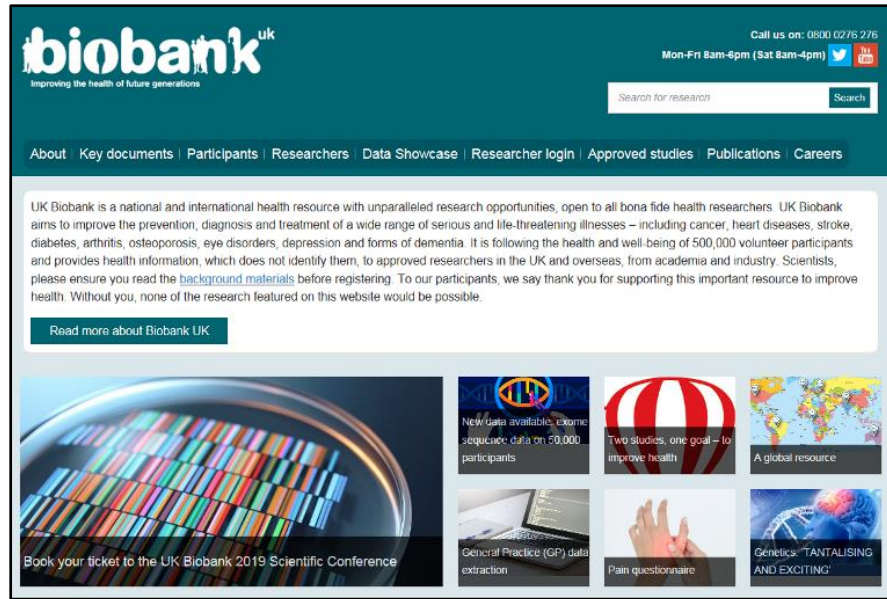
発表内容

- **大規模疾患ゲノム解析の現状**
- **ゲノム個別化医療の社会実装**
- **ゲノム創薬の実現に向けて**
- **遺伝統計学の今後の国際戦略**

ヒトゲノム研究の大規模化と感受性遺伝子同定

UKバイオバンク

(<https://www.ukbiobank.ac.uk/>)



ゲノムワイド関連解析(GWAS)

が同定した感受性遺伝子

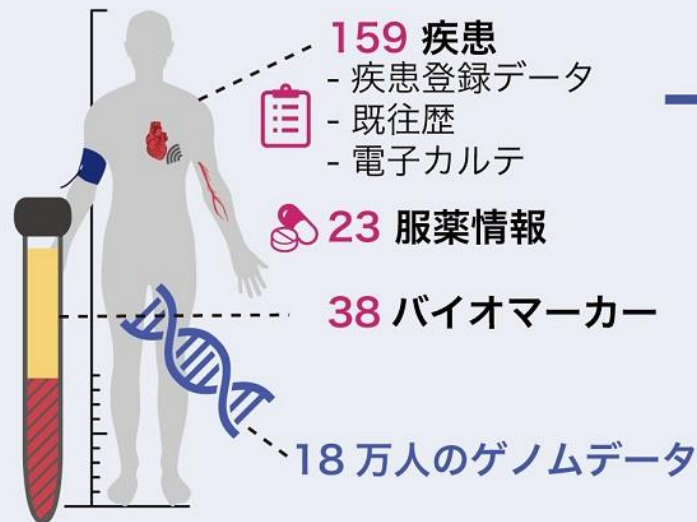
(<https://www.ebi.ac.uk/gwas/>)



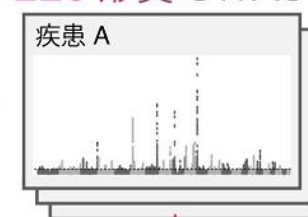
- UKバイオバンクは50万人のSNPデータ、45万人の全エクソーム、15万人の全ゲノムシーケンスを実施し、ゲノム・表現型情報を世界中に安価で分譲することで研究成果を創出している。
- 世界中の研究施設でゲノムワイド関連解析が実施され、2021年までに、**1,000以上のヒト形質**に対し、**5,000以上の論文**が報告されている。

国際バイオバンク連携によるGWASメタ解析

バイオバンク・ジャパン



220 形質 GWAS



国際バイオバンク連携

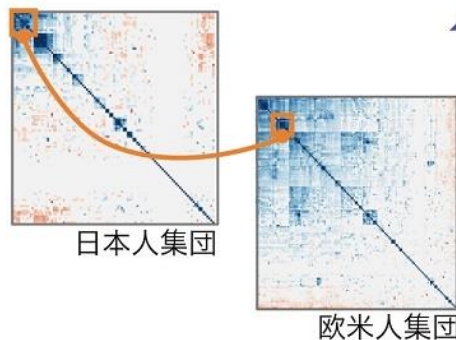
- UK バイオバンク (36 万人)
- FinnGen (14 万人)

GWAS メタアナリシス

計 63 万人
14,814 の関連領域を報告
(うち 5,343 が新規報告)

オープン
データシェア

PheWeb.jp



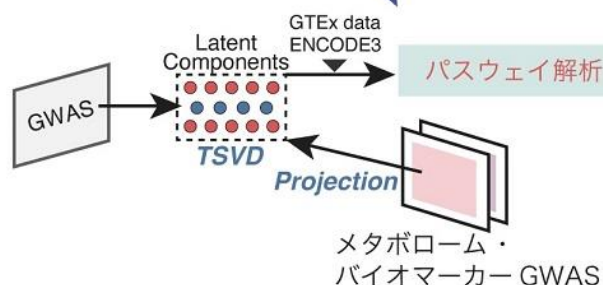
日本人集団・欧米人集団間での
遺伝的リスク構造比較

6 番染色体

MHC 領域

DPB1
DPA1
DOB1
DQA1
DRB1
DRB3
B
C
A

HLA ファインマッピング



要約統計量の特異値分解による
疾患同士のつながりの理解と再整理

・国際バイオバンク連携63万人・220形質のGWASメタ解析を実施。

・GWAS統計量を <https://pheweb.jp/> にて一般公開中。

国際バイオバンク連携によるGWASメタ解析



検索Box(形質・遺伝子・SNP)

BioBank Japan PheWeb (PheWeb.jp)

Search for a variant, gene, or phenotype

Example queries: Type 2 diabetes (T2D) PCSK9 rs671

検索例

This website releases genome-wide association study (GWAS) summary statistics of the BioBank Japan Project (BBJ). We provide GWAS results in the Japanese population (mainly from BBJ) using the PheWeb platform, with public access to the full summary statistics.

BBJ is a prospective genome biobank that collaboratively collected DNA and serum samples from 12 medical institutions in Japan, managed by the Institute of Medical Science, the University of Tokyo. BBJ has recruited approximately 260,000 participants, mainly of Japanese ancestry. All study participants had been diagnosed with one or more of 47 target diseases. RIKEN Center for Integrative Medical Sciences contributed to genotyping of the BBJ samples.

Release History

March 12, 2021: First Release!

We added 220 GWAS results from Sakaue, S. & Kanai, M., et al. (2020), Sakaue, S., et al. (2021), and Okada, Y., et al. (2014).

PheWeb.jp

<https://pheweb.jp/>

・バイオバンク・ジャパン(BBJ)を中心に、日本人集団における多彩な形質のGWASの全SNPの結果がダウンロードできるサイトです。

大規模疾患ゲノム研究のこれまでとこれから

冒険・挑戦



公共事業



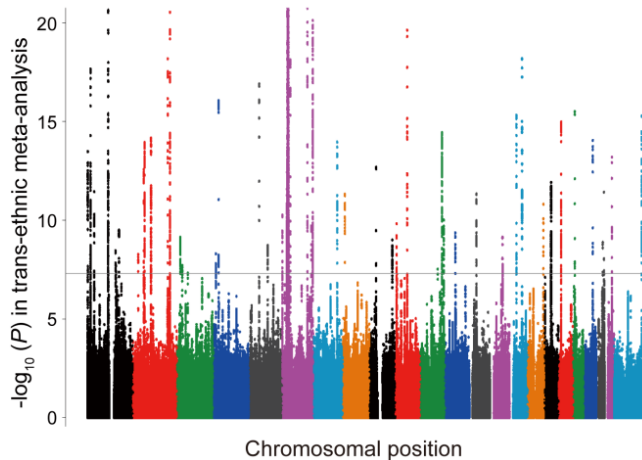
・誰もが参加できる、データに触れる、皆の研究に活用される、**サイエンスの公共事業**としての大規模疾患ゲノム研究の推進が重要。

発表内容

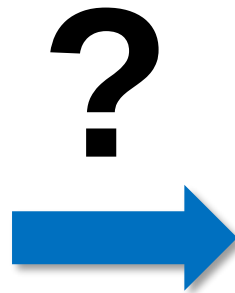
- 大規模疾患ゲノム解析の現状
- ゲノム個別化医療の社会実装
- ゲノム創薬の実現に向けて
- 遺伝統計学の今後の国際戦略

ゲノム解析の成果をどのように 活用すれば病態解明・創薬・個別化医療に 貢献できるのだろうか？

Human Genetics

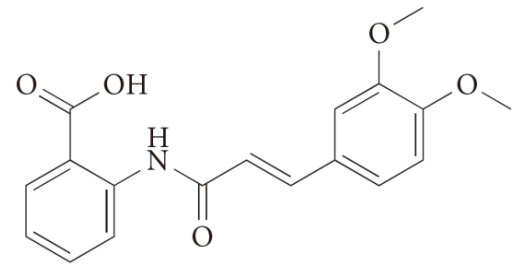


Trans-ethnic GWAS of rheumatoid arthritis identified 101 risk loci
(Okada Y. *Nature* 2014)



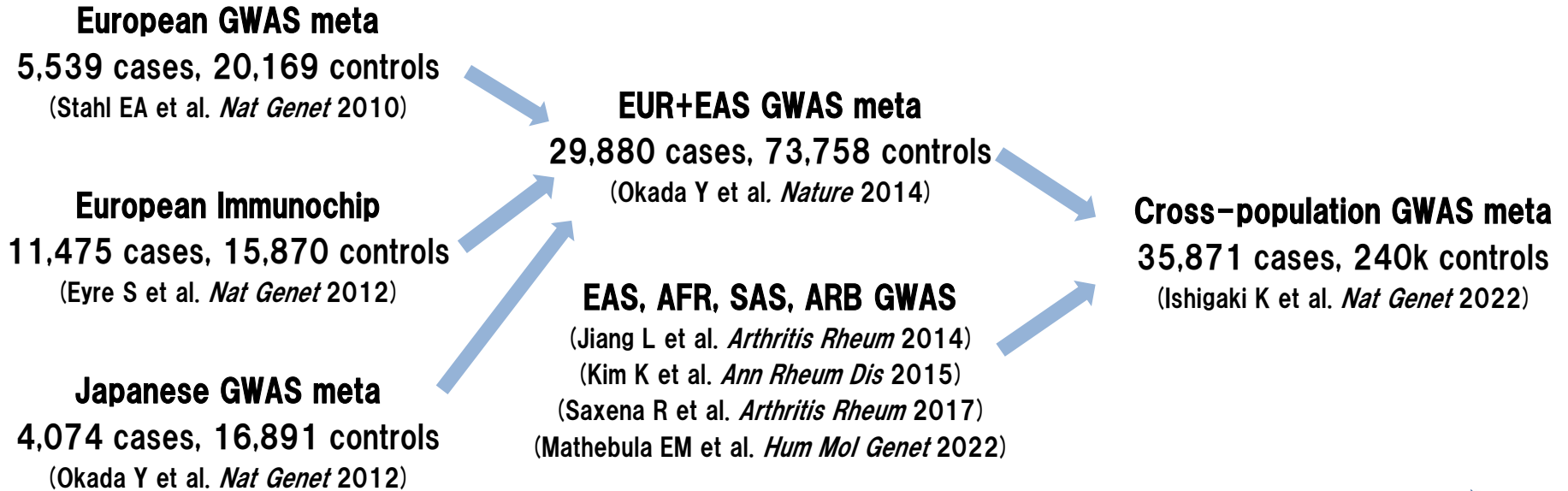
Disease
biology

Drug
discovery



Personalized
Medicine

関節リウマチGWAS解析における研究トピックの変遷



No. RA risk loci detected by GWAS

>50

>100

>150

GWASメタアナリシス
の実施方法

複数人種集団GWAS
メタアナリシスの実施方法

複数人種間で共有される
遺伝的背景の定量化

GWAS imputation
の精度評価

リスクSNPとエピゲノム
修飾のオーバーラップ

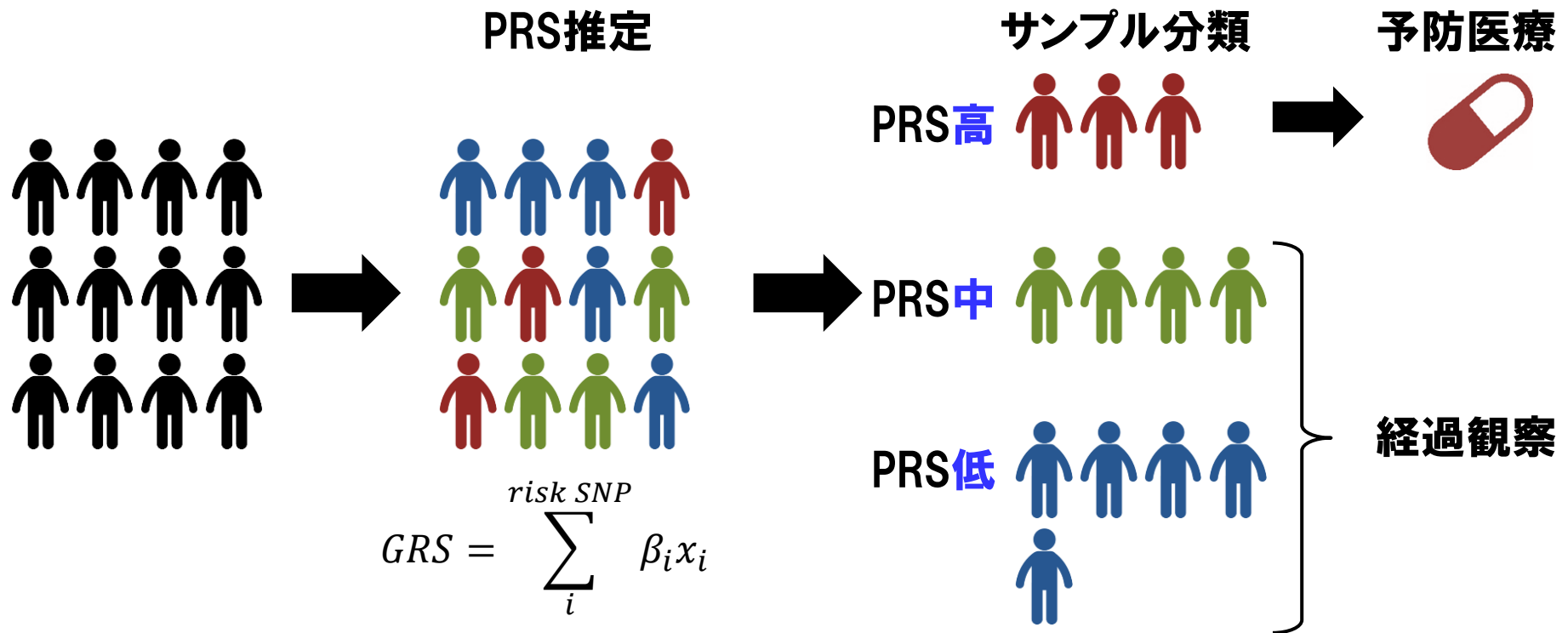
原因感受性変異の
領域内絞り込み

人種集団間で共有
される疾患感受性変異

ゲノム創薬

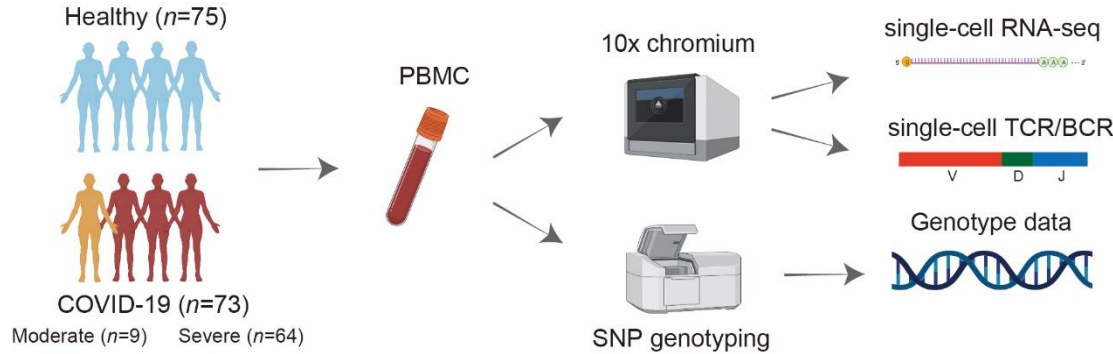
関連解析から
疾患発症予測へ¹³

Polygenic risk scoreとゲノム個別化医療

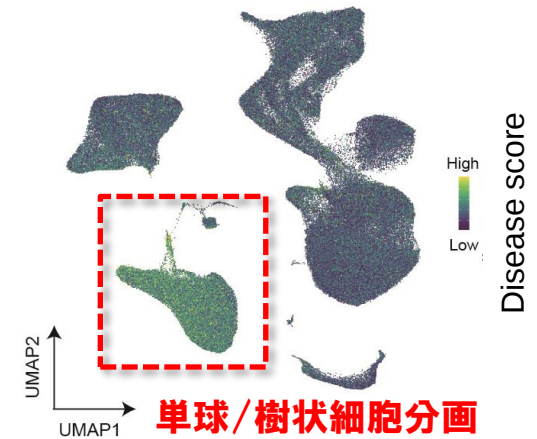


- ・生活習慣病など、集団中の一定の割合のサンプルが、加齢に伴い発症する疾患においては、予め発症リスクの高いサンプル群を同定し、**予防医療**を施すことが有用です。
- ・既知のリスク因子にPRSを加味することで、高リスク群サンプルの効率的な同定が可能になり、予防医療が推進されると期待されています。

日本人集団COVID-19患者血液シングルセル解析



● Polygenic risk of severe COVID-19

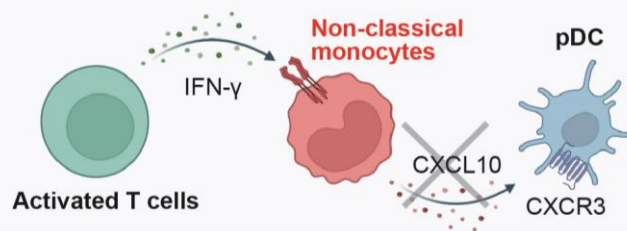


Non-classical monocytes in severe COVID-19



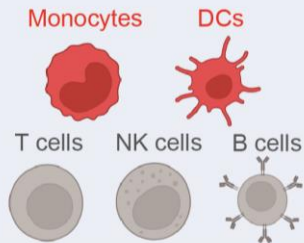
- Cell proportion ↓
- Cell transition from classical monocytes ↓
- Responsset to IFN ↓
- Cellular interactions ↓

CXCL10/CXCR3 interaction in severe COVID-19



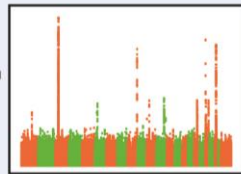
- IFN- γ signal from activated T cells to non-classical monocyte ↑
- CXCL10 ↓ in non-classical monocytes
- CXCL10/CXCR3 ↓ between non-classical monocytes and pDC

Cell type enrichment of host genetics COVID-19 risk

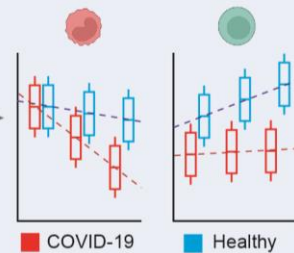


- Enrichment of Monocytes and DCs in severe phenotype

COVID-19 GWAS

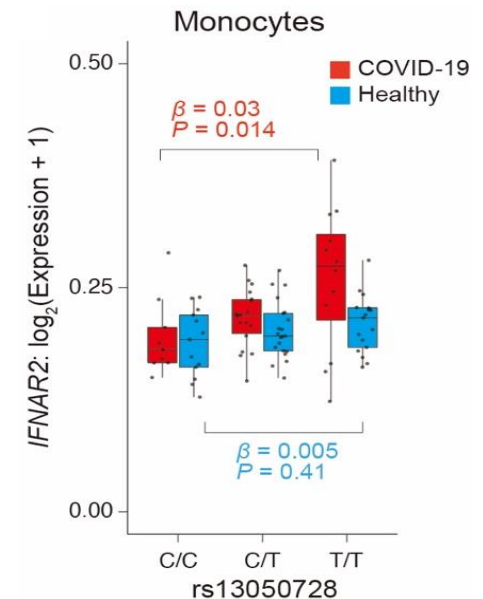


eQTL analysis



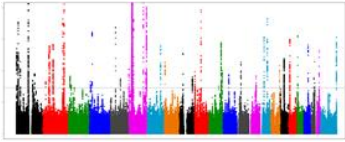
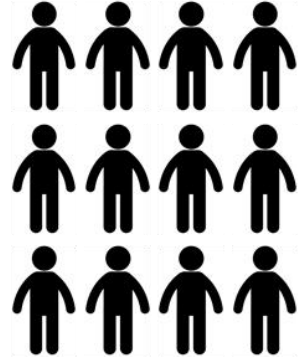
- Context & cell type-specific eQTL effect

● Context-specific eQTL of COVID-19 risk variants (e.g., *IFNAR2*)



オミクス・臨床情報に基づく集団層別化・ゲノム個別化予防

大規模ゲノム
データ解析

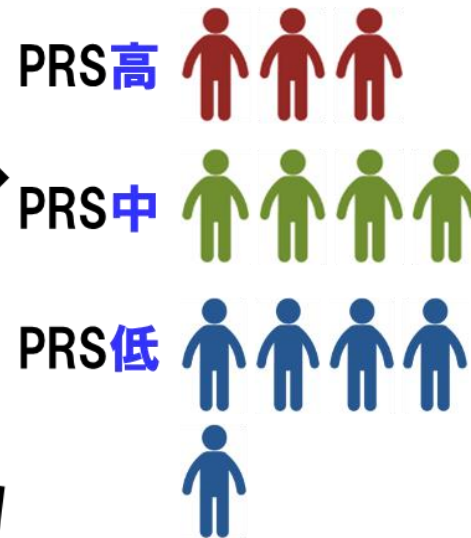


Polygenic Risk
Score推定



$$PRS = \sum_i^{risk\ SNP} \beta_i x_i$$

集団サンプル
層別化モデル



ゲノム
個別化予防

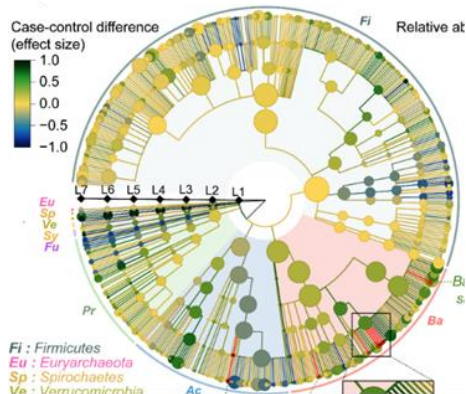


標準的
観察手段

ゲノム個別化医療
の社会実装にむけて
・今後の課題設定

前向きコホート・バイオバンク由来の情報・臨床情報

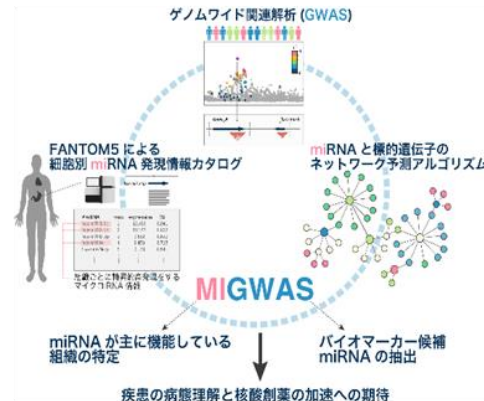
腸内細菌メタゲノム



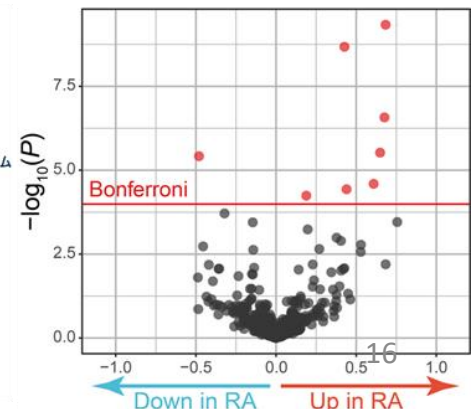
臨床情報・生活習慣

- ・臨床情報
- ・頭部画像情報
- ・既往歴
- ・生活習慣
- ・食生活習慣
- ・環境因子
- ・ライフログ

組織別エピゲノム



メタボローム



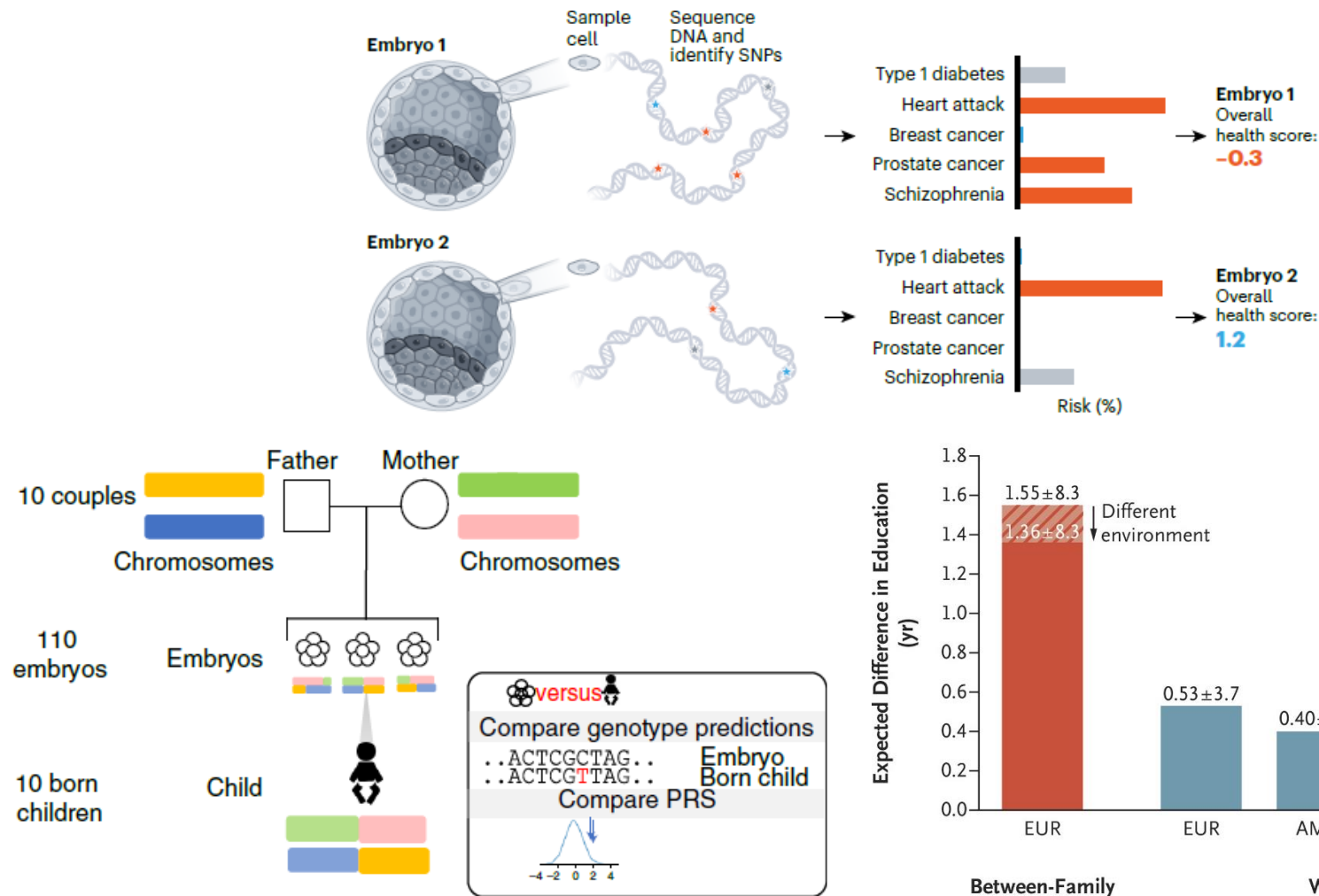
Polygenic risk scoreとゲノム個別化医療

PRSを個人に還元するまでの「遠い」道のり



- PRSによる疾患発症予測精度向上のための課題整備が進められている。
- 一方、「**得られたPRSをどのように個人に還元するか**」については、未だ具体的な道筋が得られていない。
- **医療機関・アカデミア・ベンチャーの密な連携**を通じてゲノム個別化医療社会実装に取り組んでいく時代が到来している。

出生前診断におけるPRS活用を巡る懸念



• 出生前診断において、複数の胚由来のゲノム情報からPRSを推定し、特定の形質における推定値に基づき胚を選択する医療行為について、懸念が提唱されています。

(Turley P et al. *NEJM* 2021, Forzano F et al. *Eur J Hum Genet* 2021, Kozlov *Nature* 2022, Kumar A et al. *Nat Med* 2022)

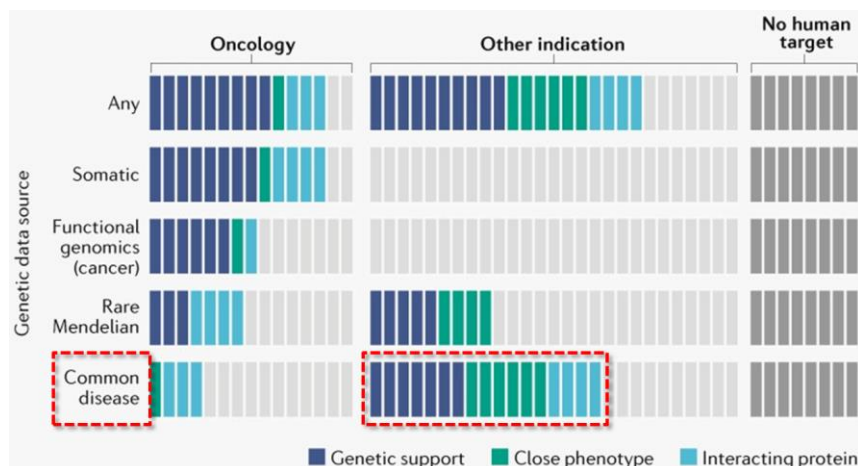
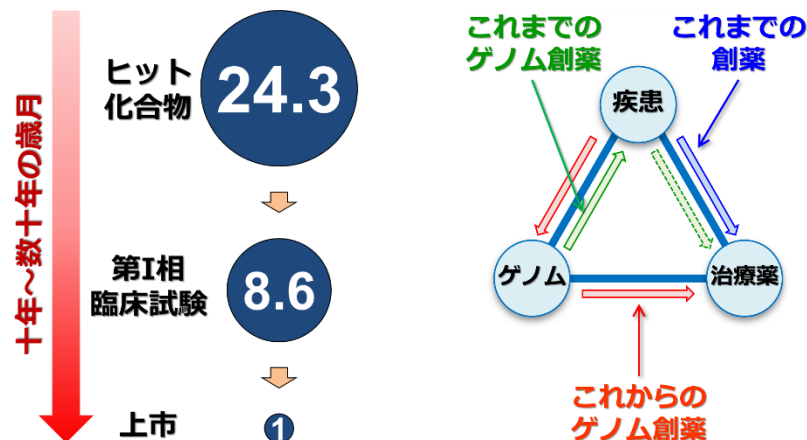
発表内容

- 大規模疾患ゲノム解析の現状
- ゲノム個別化医療の社会実装
- ゲノム創薬の実現に向けて
- 遺伝統計学の今後の国際戦略

ゲノム創薬による創薬プロセスの効率化

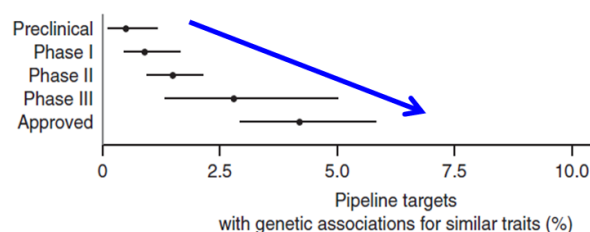
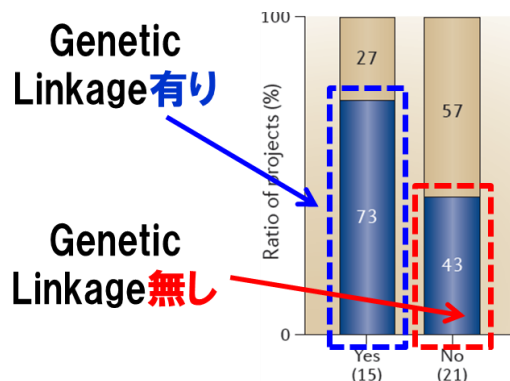
創薬コスト増大・これからのゲノム創薬

FDA承認薬の2/3は疾患ゲノム情報にサポートされる

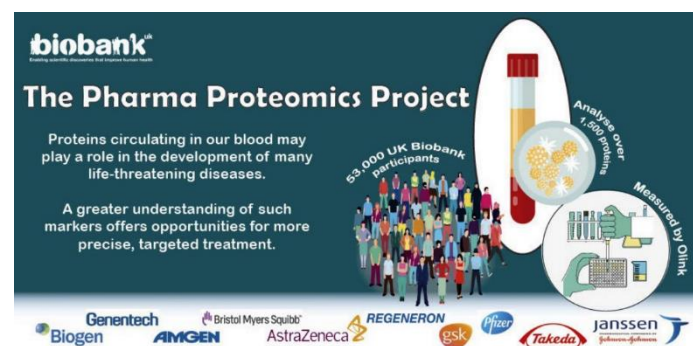


海外製薬企業におけるゲノム創薬導入と創薬プロセス効率化

企業主導型UK Biobankプロテオーム解析



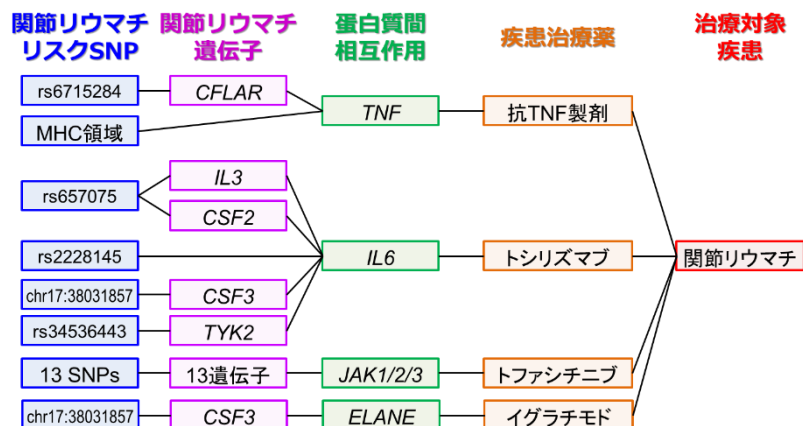
創薬ステージが進むにつれ
Genetic Evidenceが増す



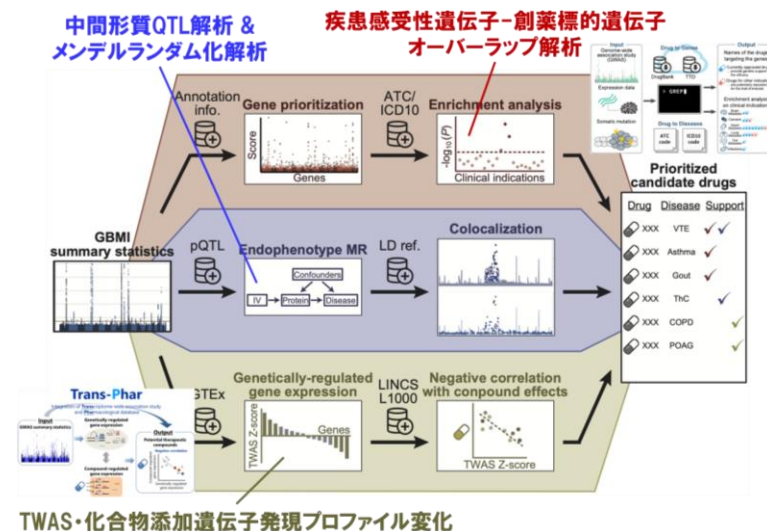
- ・創薬コスト増大・成功率低下の打破には、疾患ゲノム情報を活用する「ゲノム創薬」が不可欠。
- ・疾患ゲノム解析を通じて構築された、「遺伝子変異⇔ヒト形質の因果関係」のカタログを活用。
- ・海外製薬企業において、ゲノム創薬が創薬プロセスを効率化することが示されている。
- ・バイオバンクにおいても、ゲノム創薬に有効な中間形質QTLデータベース構築が進んでいる。

ゲノム創薬研究の実例

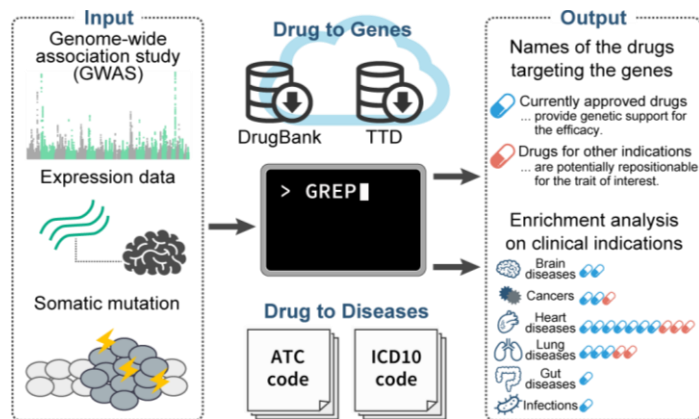
関節リウマチGWASと創薬シーズのつながり



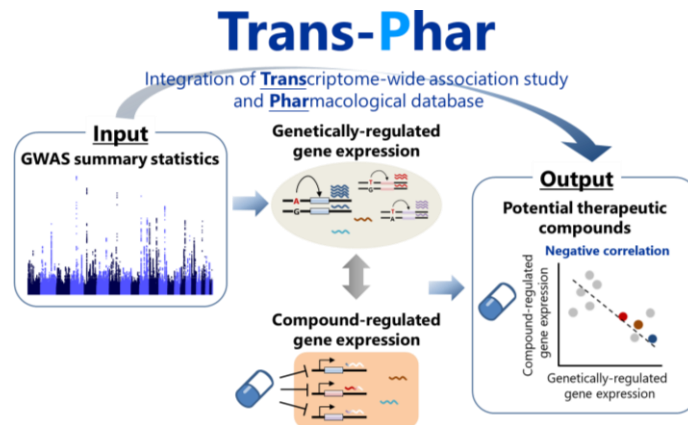
ゲノム創薬ベストプラクティスの国際提言



GREP: WHO ATC分類に基づく自動リポジショニング

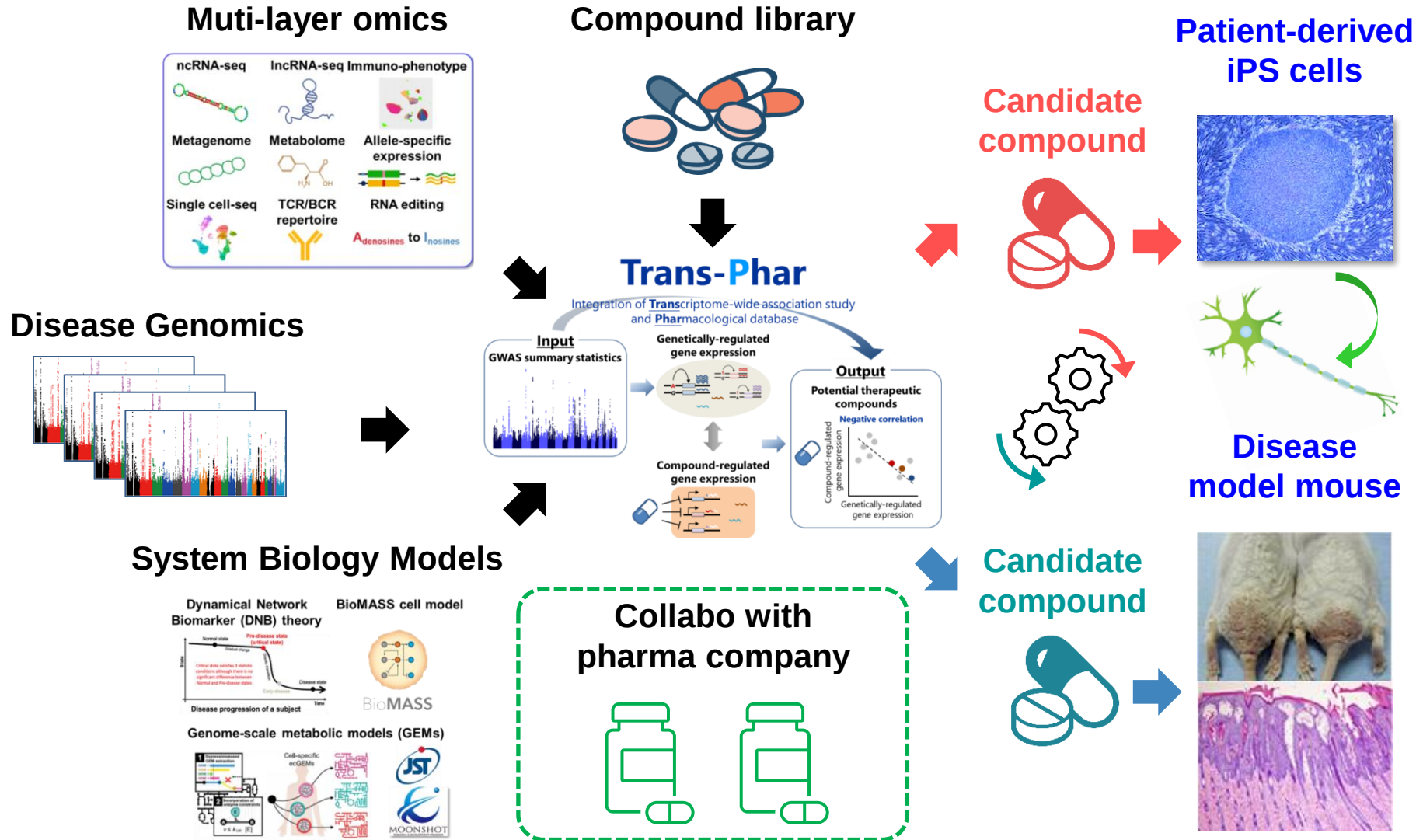


Trans-Phar: 遺伝子発現変化ダイナミクスを反映



- GWASなど疾患ゲノム情報に基づくリポジショニング化合物スクリーニングのアルゴリズムを実装。
- 関節リウマチ、脳梗塞など、疾患ゲノム情報に基づく候補化合物同定の実績。

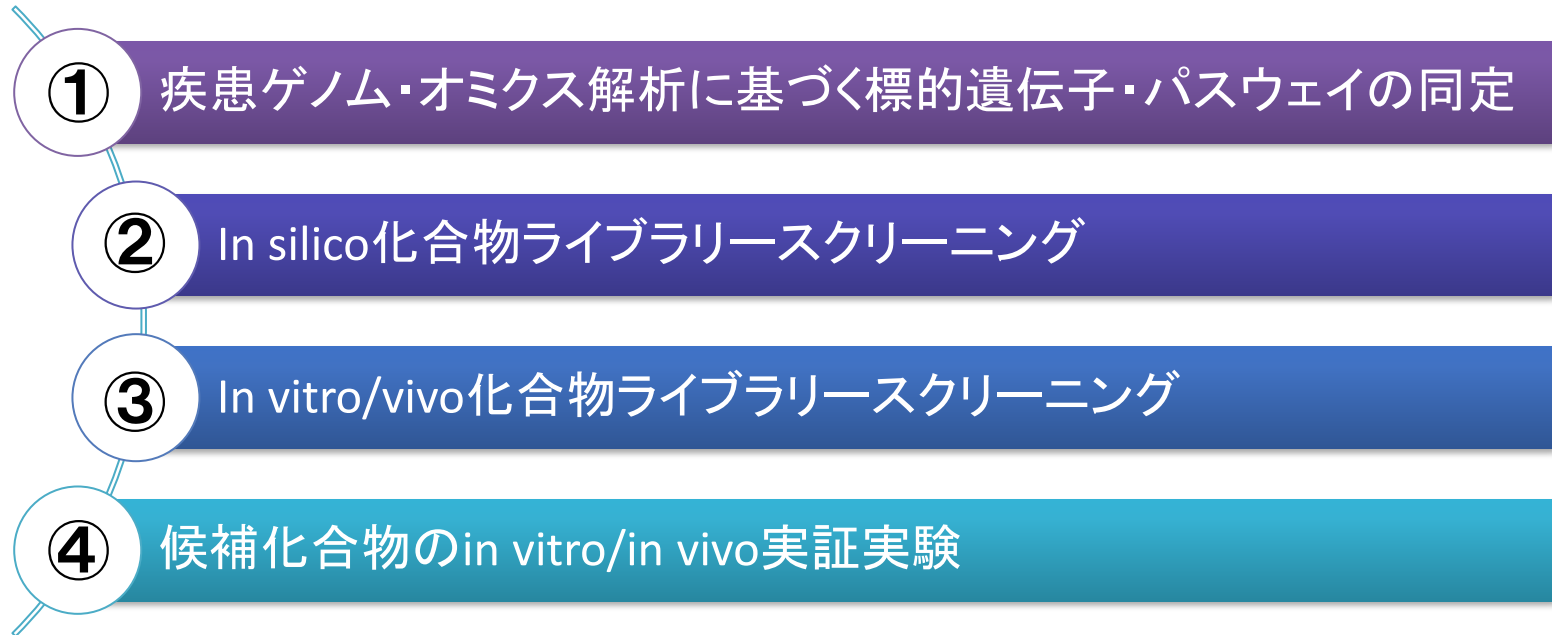
これからのゲノム創薬 (in silico → in vivo → at bed)



- ・対象疾患のゲノミクス解析結果より創薬標的をスクリーニング。
- ・患者由来iPS細胞への化合物添加を実施し、患者層別化と共に評価。

これからのゲノム創薬実装の方向性

ゲノム創薬実装における段階的プロセス



- ・ゲノム創薬実装プロセスでは、① ⇒ ② ⇒ ③ ⇒ ④のシームレスな連携が不可欠。
- ・② ⇒ ③、③ ⇒ ④を重点的につなぐことで、異分野連携を軸とした実装が期待される。

- ①: 標的遺伝子・パスウェイ同定 ⇒ 遺伝統計学を軸としたゲノム・オミクス解析で実装済
- ②: In silicoスクリーニング ⇒ ゲノム創薬手法でカバー可能
- ③: In vitro/in vivoスクリーニング ⇒ ハイスループットな系が望ましい。異分野連携を模索。
- ④: 実証実験 ⇒ iPS分化細胞やオルガノイド実験系、前向きコホートで実施可能。

発表内容

- 大規模疾患ゲノム解析の現状
- ゲノム個別化医療の社会実装
- ゲノム創薬の実現に向けて
- 遺伝統計学の今後の国際戦略

遺伝統計学分野の日本の実情

- ・過去20年間、アジア人集団におけるゲノムデータの貴重性に価値が見出されていた。
- ・世界各国がバイオバンクを有する現在、新たな舵取りが求められている。
- ・国際コンソーシアム運営、国際ガイドライン策定における存在感の低下。
- ・欧米との解析技術格差が開きつつある。ゲノムシーケンスは実施規模に乖離。オミクス解析技術は技術開発から現場導入までのタイムラグが延長。
- ・導入後の解析機器を器用に使いこなすノウハウは長けている。
- ・遺伝統計学の専門家の若手人材不足解消が急務。
- ・医学部出身者のデータ解析研究者は少しずつ増えてきている印象。
- ・医学-オミクス解析の融合研究は成果が出ている。
- ・数理科学、情報科学、統計学の専門家の参入には改善の余地がある。
- ・ゲノム・オミクスデータの公開・共有・非属人化の遅れ。
→ 最先端のデータに触れる機会の喪失が参入障壁となり、更なる人材不足を招く。

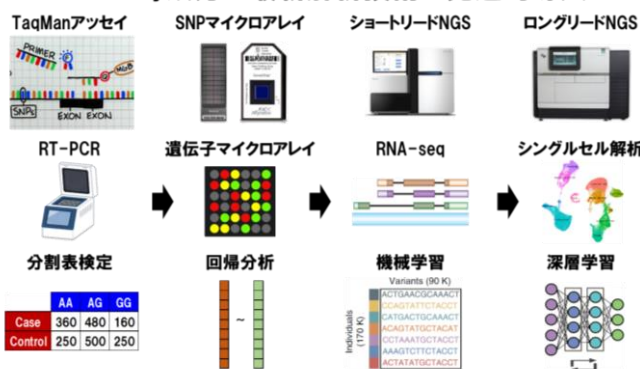
グローバルバイオバンク時代の戦略

- 1. Aim:** Construction of genotype-phenotype catalogue for health promotion.
Method: To GWAS SNP array genotype all the samples.
- 2. Aim:** Population-specific rare functional variant catalogue for drug discovery.
Method: To WGS sequence the valuable samples (rather than WES).
- 3. Aim:** Construction of trans-omics resources for disease biology.
Method 1: Omics from serum (NMR metabolome, oligo proteome).
Method 2: Omics from cells (blood & organ single cell RNA-seq).
Method 3: Omics from clinical EMR and images (e.g., lung CT, brain MRI).
- 4. Aim:** Continuous enhancement of biobank resources.
Method: Bi-directional collection of participants and clinical info updates.
- 5. Aim:** International initiative in the era of “biobank**s** in every country”.
Method 1: Promotion of cross-biobank collaboration.
Method 2: Participants into global biobank network.

世界に勝つための戦略

- ・医学科・生命科学科における、キャリア早期のバイオインフォマティクス演習の導入。
- ・若手研究者の海外留学(ポスドク・博士課程)の積極奨励(制度・財政面)。
- ・「アジアの雄」を「真剣に」目指す。英語化推進による国際競争力の底上げ。
- ・Wet・Dry解析機器のコアファシリティ化による、最新解析技術の積極導入。共同研究ではなく業務としてのオミクス解析・知財出願のサポート。関連人材のキャリア形成。
- ・数理科学、情報科学、統計学分野の専門家へのアピール(=待遇面の改善も)。
- ・Curiosity, Methodology, Mission-driven Initiativeのバランスの再認識。
- ・大規模ゲノム・バイオバンクデータの公開・共有・非属人化促進。

Wet・Dry双方の最新解析技術の先進的導入



Methodology-driven イニシアティブ



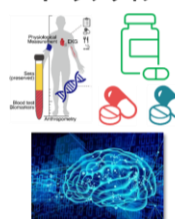
We have more Data than we can analyze.

Curiosity-driven イニシアティブ



Why want to climb Everest? - Because it is there.

Mission-oriented イニシアティブ



Hippocratic Oath

- ・Curiosity: 基礎研究の根幹を成す、生命現象解明への探求心。
- ・Methodology: 最新解析技術が可能にする、生命科学の新展開。
- ・Mission: 医学研究として不可欠な、疾病克服・健康長寿への挑戦。
- ・3者の融合を通じて、ライフサイエンスのイニシアティブを探索していく。

遺伝統計学
ONLINE
夏
の
学
校

夏休みの思い出に、ゲノムの世界に触れてみませんか？

日時：令和5年8月25日（金）～8月27日（日）
（原則として三日間通じて参加をお願いします。ご都合に応じて一部日程での参加も受け付けておりますので、事務局までお問合せ下さい）

場所：オンライン形式でのWeb開催
（複数研究施設による共同開催の場合、オンライン形式での開催になります）

対象：大学学部生以上
（大学院生・社会人の方も歓迎です。保護者の了解が得られれば高校生も歓迎です）

形式：講義、ゲノム・オミクスデータ解析演習
（演習に使用するパソコンは、参加者各自で用意して頂く必要があります）
（使用する解析用データ形式は、夏の学校で用意し、設定資料と共に事前配布します）

定員：100名
（応募者多数の場合、所属や学年を考慮して調整させて頂く事があります）

日程表

時間帯	一日目 (8/25)	二日目 (8/26)	三日目 (8/27)
午前 (9:00-12:00)	遺伝学入門	統計ソフト入門	ゲノム・オミクスデータ解析演習①～③
午後 (13:00-16:30)	Linux入門	ゲノム・オミクスデータ解析演習④～⑥	

（実際の講義演習内容は、変更となる場合があります）

URL : http://www.sg.med.osaka-u.ac.jp/school_2023.html
申込み : office@sg.med.osaka-u.ac.jp (担当：岡田 隆象)

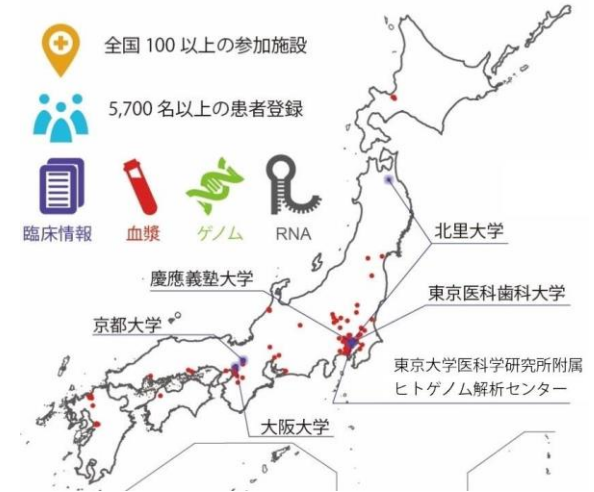
主催：大阪大学大学院医学系研究科遺伝統計学
共催：東京大学大学院医学系研究科遺伝統計学、理化学研究所生命医科学研究センターシステム遺伝学チーム
後援：大阪大学バイオインフォマティクス・イニシアティブ、会費ムーンショットプロジェクト

- ・観測・解析技術の進歩は、常に予想を上回る速度で進む。
- ・Wet・Dry双方の最新解析技術の先進的導入が生命科学に不可欠。
- ・実験・解析原理の正確な理解が、革新的な研究を可能にする。

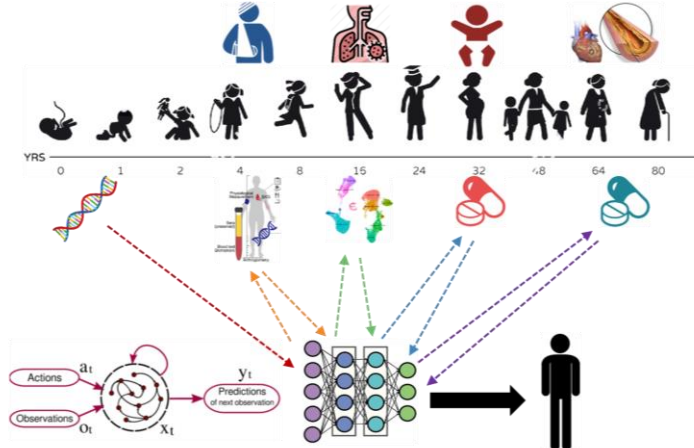
遺伝統計学の今後の展望

- 誰でも世界最高峰のデータにアクセスできる時代、に即した戦い方が必要。
- 医学、情報学、数理科学、生物学、疫学、多彩なバックグラウンドを持つ専門家チームが、個の技と総合力で勝負していく。
- 疾病の解明 → 予測 → 制御、を科学する学問へ変革。
- ライフイベントの揺らぎを、経時的に、制御。
- 網羅的Dry解析 ⇔ In vivo実証実験、の逆還元。
- データ解析研究のコスト認識に配慮した予算配分。
- 「楽しく奥深い基礎研究」としての矜持。

コロナ制圧タスクフォース

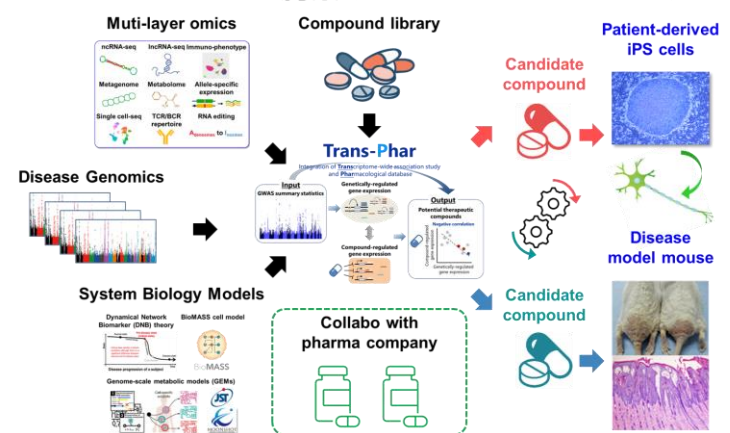


ヒトゲノム情報への機械学習応用と個別化医療



「Uni-directional」から「bi-directional & feedback-looping」機械学習への転換による「ヒト人生」の予測。

これからのゲノム創薬 (in silico → in vivo → at bed)



- 対象疾患のゲノミクス解析結果より創薬標的をスクリーニング。
- 患者由来iPS細胞への化合物添加を実施し、患者層別化と共に評価。

次代を担う若手研究者の人材育成

遺伝統計学・夏の学校の開催
(2023年度はオンライン開催:8/25-27)

医学部正規カリキュラムへの
ゲノムデータ解析演習導入

遺伝統計学 ONLINE
参加費: 無料

夏の学校

夏休みの思い出に、ゲノムの世界に触れてみませんか?

日時: 令和5年8月25日(金)~8月27日(日)
(原則として三日間通じて参加をお願いしています。ご都合に応じて一部日程での参加も受け付けておりますので、事務局までお問合せ下さい)

場所: オンライン形式でのWeb開催
(複数研究施設による共同開催の都合、オンライン形式での開催になります)

対象: 大学学部生以上
(大学院生・社会人の方も歓迎です。保護者の了解が得られれば高校生も歓迎です)

形式: 講義、ゲノム・オミクスデータ解析演習
(演習に使用するパソコンは、参加者各自で用意して頂く必要があります)
(使用する解析用データ一式は、夏の学校で用意し、設定資料と共に事前配布します)

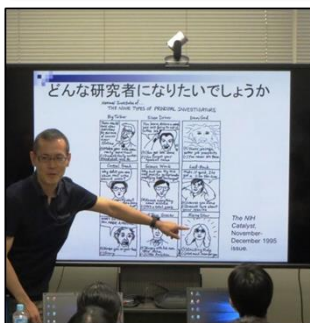
定員: 100名 (応募者多数の場合、所属や学年を考慮して調整させて頂くことがあります)

時間帯	一日目 (8/25)	二日目 (8/26)	三日目 (8/27)
午前 (9:00~12:00)	遺伝学入門 統計学入門	統計ソフトR入門	ゲノム・オミクス データ解析演習 ④~⑦
午後 (13:00~16:30)	Linux入門 プログラミング入門	ゲノム・オミクス データ解析演習 ①~③	

(実際の講義演習内容は、変更となる場合があります)

URL : http://www.sg.med.osaka-u.ac.jp/school_2023.html
申込み: office@sg.med.osaka-u.ac.jp (担当: 岡田 随象)

主催: 大阪大学大学院医学系研究科遺伝統計学
共催: 東京大学大学院医学系研究科遺伝統計学、理化学研究所生命医科学研究センターシステム遺伝学チーム
後援: 大阪大学バイオインフォマティクス・イニシアティブ、合衆ムーンショットプロジェクト



③ Cygwinを使ったLinuxコマンド実習		
コマンド	機能	
top	システム使用状況を表示	
pwd	カレントディレクトリを表示	
cd	カレントディレクトリを変更	
ls	ディレクトリやファイル情報を表示	
echo	文字列や変数の値を標準出力に表示	
cp	ファイルやディレクトリをコピーする	
mv	ファイルやディレクトリを移動する	
rm	ファイルやディレクトリを消去する	
wc	テキストファイルの大きさや行数を表示	
gzip	ファイルとgz形式で圧縮	
gunzip	gz形式のファイルを解凍	
less	ファイル内容を画面ごとに表示	
head	ファイル内容を行単位で先頭から表示	
tail	ファイル内容を行単位で末尾から表示	

・Linuxコマンドは多数存在します。頻回に使用するものから覚えましょう。

- 公開データを用いたゲノム解析体験の提供を通じて、有望な国内若手人材を発掘中。
- ゲノムデータ解析実践演習講義「**遺伝統計学・夏の学校**」を毎年自主開催し、日本各地から若手バイオインフォマティシャン(の卵)が参加。羊土社より書籍化。
- ゲノムデータ解析演習を、**大阪大学・東京大学医学部の実習カリキュラム**として開始。