

遺伝統計学の 国際展開戦略

岡田 随象

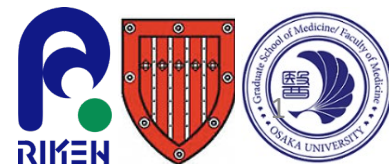
大阪大学大学院医学系研究科 遺伝統計学

東京大学大学院医学系研究科 遺伝情報学

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 : yokada@sg.med.osaka-u.ac.jp



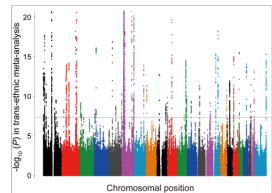
遺伝情報と形質情報の関わりを検討する**遺伝統計学**を主軸に、多彩なヒトオミクス情報のデータ解析を通じて、**疾患病態解明・ゲノム創薬・個別化医療**の実現に取り組んでいます。

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

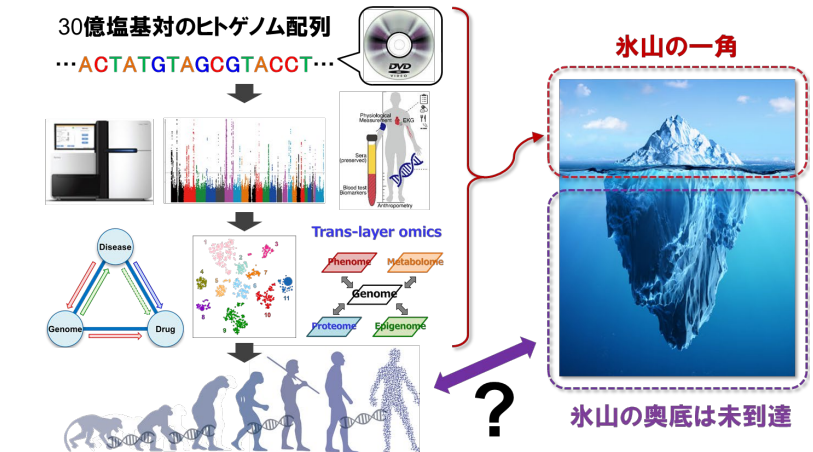
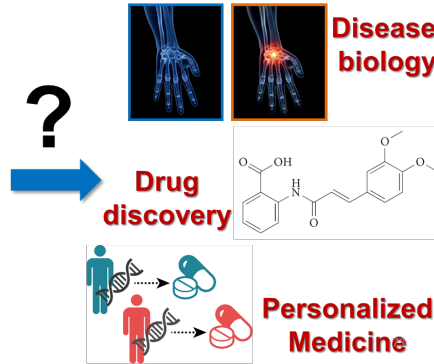
ゲノム解析の成果をどのように活用すれば
病態解明・創薬・個別化医療に貢献できるのか
ほとんどわかっていない

氷山の一角: ヒトゲノム配列のほとんどは未解明

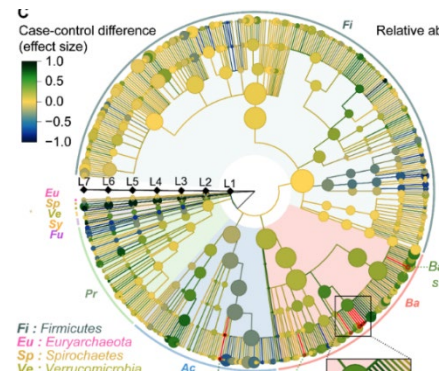
Human Genetics



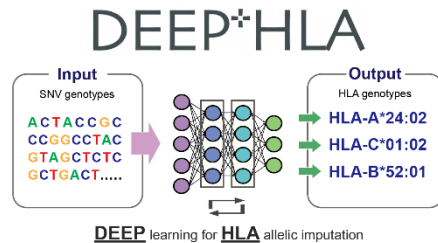
Trans-ethnic GWAS of rheumatoid arthritis identified 101 risk loci (Okada Y. *Nature* 2014)



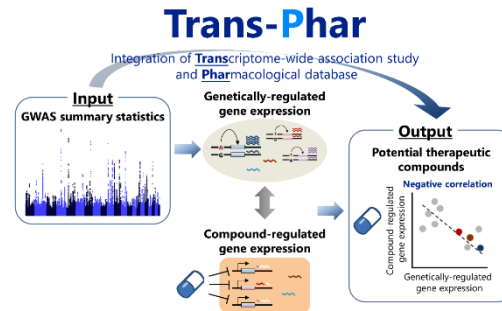
**腸内微生物叢メタゲノム
ショットガンシーケンス**



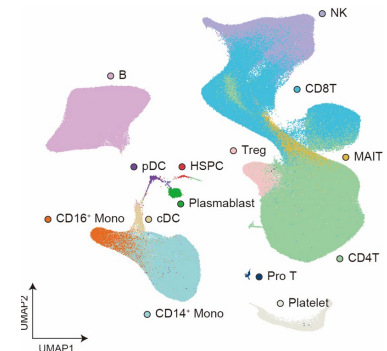
**集団ゲノム情報への
畳み込み深層学習
(DEEP*HLA)**



**疾患ゲノム創薬の実装
(Trans-Phar)**

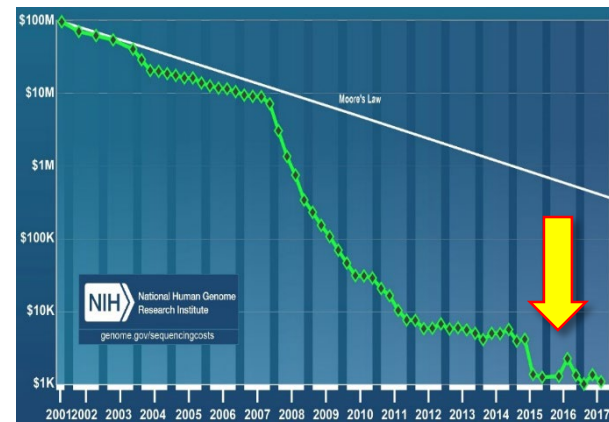


**シングルセル解析
による疾患病態解明**

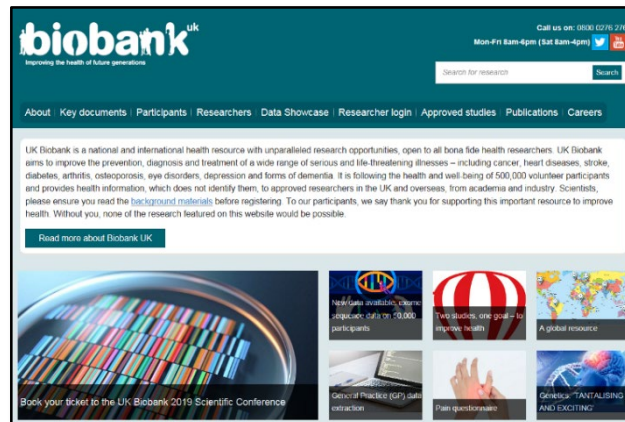


ゲノム配列情報取得コストの劇的な低下

1サンプルの全ゲノムシーケンスコスト



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

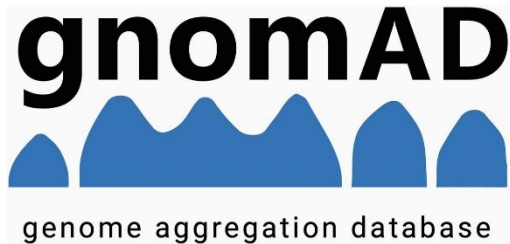


- ヒト全ゲノム配列情報の取得コストは劇的に低下し、\$1,000/サンプルが達成され、\$100/サンプルも視野に入るなど、ヒトゲノム研究への参入障壁が大幅に低下。
- SNPマイクロアレイ情報の取得コストは、\$40/サンプルが達成されている。
- UKバイオバンクは50万人のSNP・全エクソームデータ、15万人の全ゲノムシーケンスを実施し、ゲノム・表現型情報を世界中に分譲することで研究成果を創出。
- 世界中でゲノムワイド関連解析が実施され、2021年までに、1,000以上のヒト形質に対し、5,000以上の論文報告。数多くの疾患感受性遺伝子変異が同定されている³。

国際コンソーシアムによるヒトゲノム研究の推進

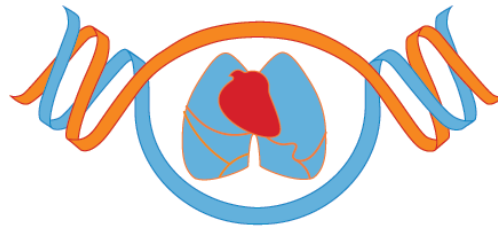
gnomADコンソーシアム

(<https://gnomad.broadinstitute.org/>)



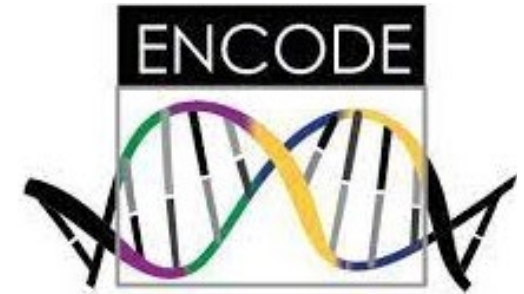
TopMedコンソーシアム

(<https://topmed.nhlbi.nih.gov/>)



ENCODEコンソーシアム

(<https://www.encodeproject.org/>)



GIANTコンソーシアム

(https://portals.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium)



Human Cell Atlas

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



GTEXコンソーシアム

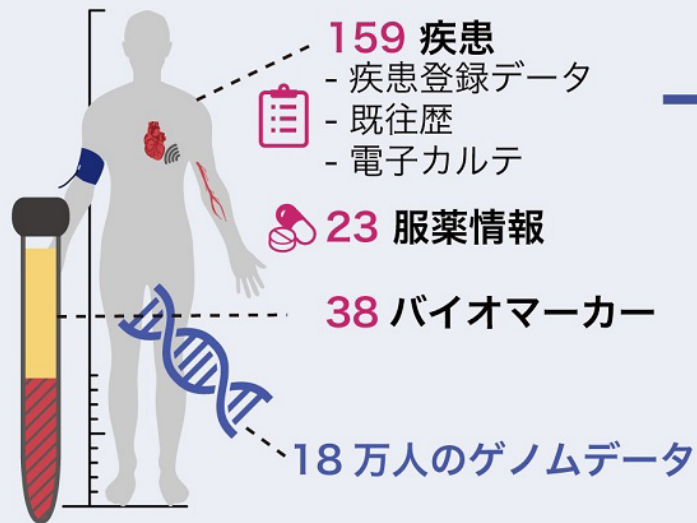
(<https://gtexportal.org/home/>)



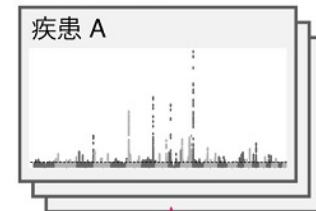
- 国際コンソーシアムを通じて、大規模ゲノミクス情報を効率的に構築、公開する時代。
- 国際イニシアティブの獲得競争が激化している。

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



GWAS メタアナリシス

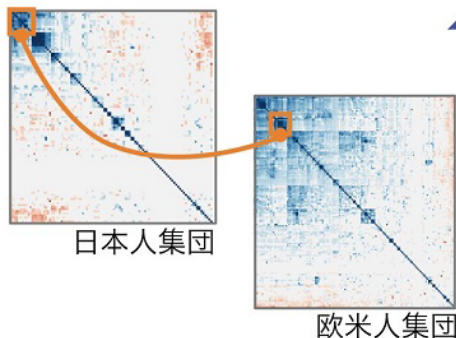
計 63 万人
14,814 の関連領域を報告
(うち 5,343 が新規報告)

国際バイオバンク連携

- UK バイオバンク (36 万人)
- FinnGen (14 万人)

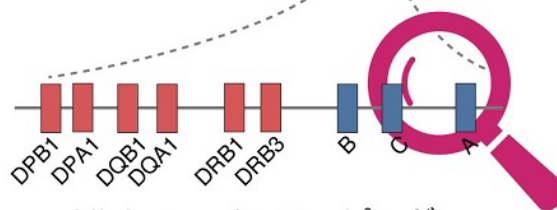
オープン
データシェア

PheWeb.jp

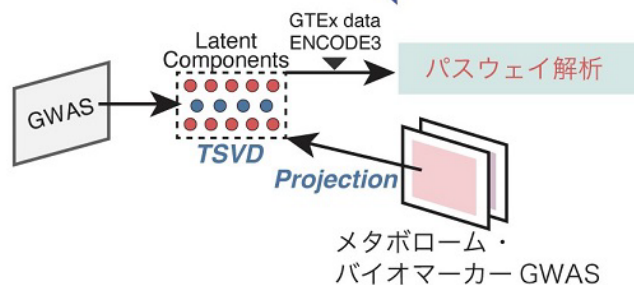


日本人集団・欧米人集団間での
遺伝的リスク構造比較

6 番染色体 MHC 領域



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(バイオバンクジャパン、UKバイオバンク、FinnGen)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。

コロナ制圧タスクフォース: オールジャパンCOVID-19研究

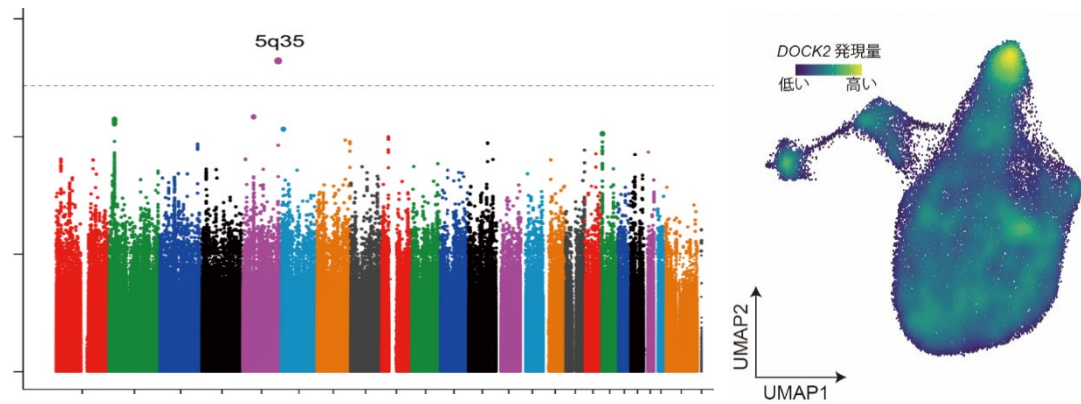
全国規模100以上の参加施設
(主催: 慶応義塾大学)

コロナ制圧タスクフォース

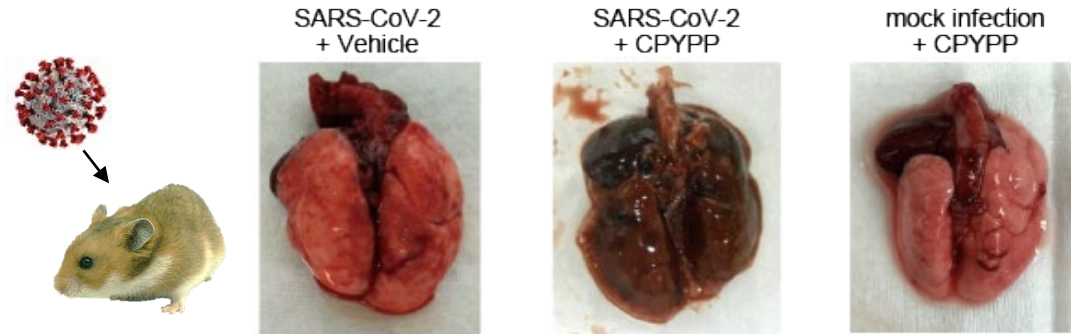


<https://www.covid19-taskforce.jp/>

GWAS・シングルセル解析: 重症化遺伝子DOCK2同定



DOCK2阻害剤によるCOVID-19重症化メカニズム解明



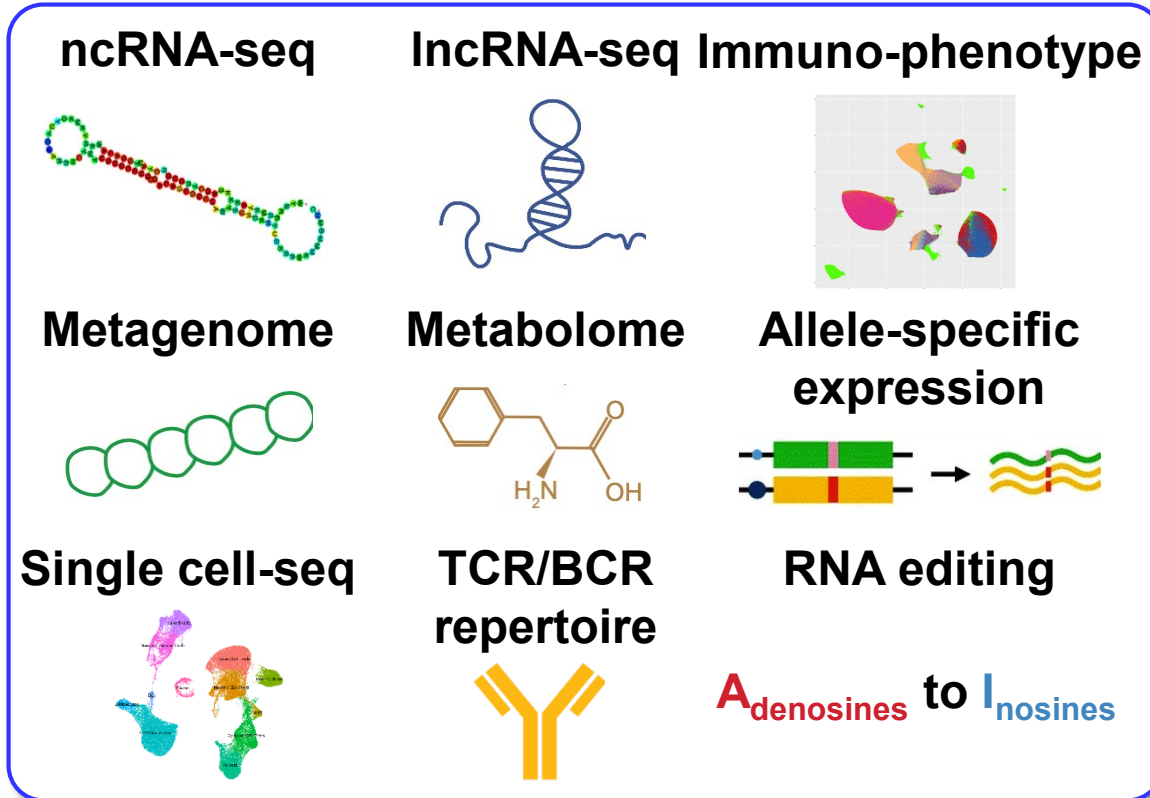
- 日本人集団における新型コロナウイルスの制圧を目指し、2020年5月にコロナ制圧タスクフォースを結成。感染者6,000名のゲノム・RNA・血漿・臨床情報を収集済。
- アジア人集団最大規模の新型コロナウイルス研究コホートとして、ゲノム・オミクス情報に基づく患者・健常者集団層別化により、効率的な予防医療(ワクチン接種)と先進医療の社会実装へ。

All Handai Project: COVID-19研究



Human genome variations

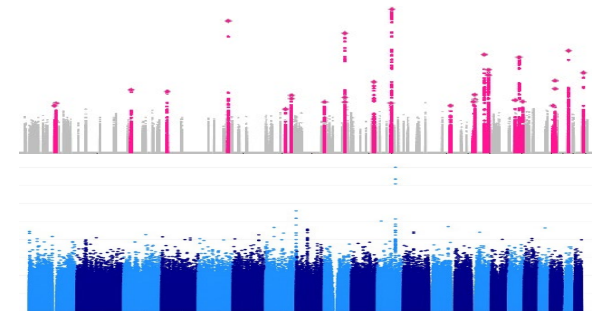
...ACTTGT **A** GCTACT...
 ...ACTTGT **C** GCTACT...



“All-Handai” Project



Tran-omics QTL database of Japanese



Cross-population database construction

• “All Handai Project”として、大阪大学の生命系部局(医・IFReC・微研)の若手研究者の連携で、**シングルセル解析**を中心にCOVID-19オミクス解析を集約・協調して実施。

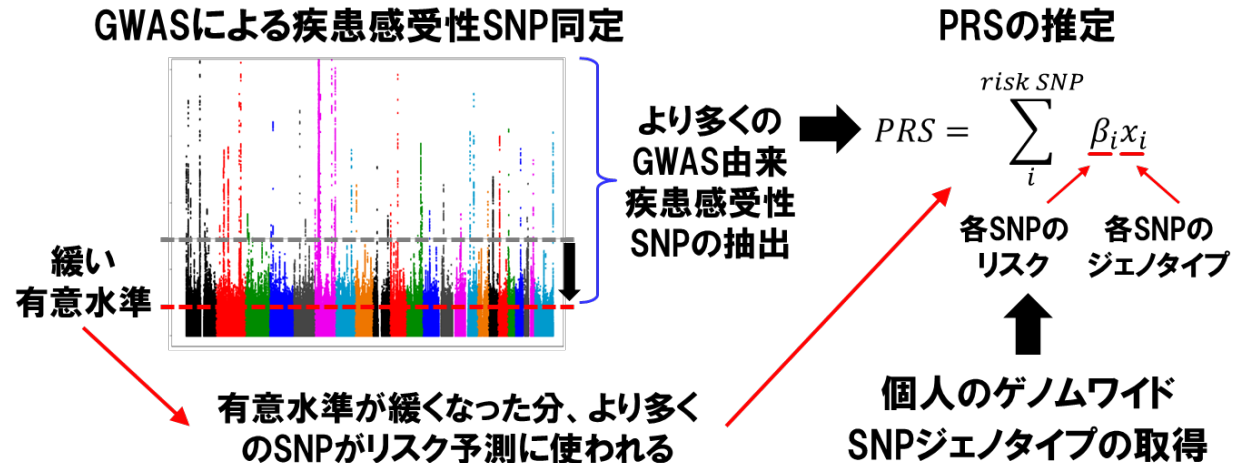
ゲノム個別化医療時代の幕開け



Precision Medicine Initiative

遺伝子、環境、ライフスタイルに関する個人ごとの違いを考慮した予防や治療法を確立すること

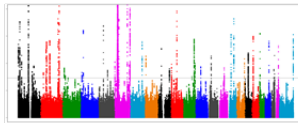
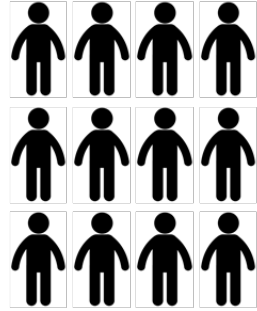
Polygenic Risk Scoreによる個人の疾患発症リスク定量化



- "Precision Medicine Initiative"は、個人のゲノム情報の違いを活用した個別化医療の提言に重きを置いた点が画期的。ゲノム個別化医療の時代の幕開け。
- 疾患GWAS結果と、個人のジェノタイプ情報を組み合わせ、個人の疾患発症リスクを定量化するPolygenic Risk Score (PRS)がゲノム個別化医療の鍵として注目。
- 個々の感受性SNPが持つリスクは小さくても、ヒトゲノム領域全体に分布する無数の疾患感受性遺伝子変異を組み合わせることで、効果的な予測が可能に。
- 欧米人集団心血管障害PRS解析では、PRS上位0.5%の集団が5倍の発症リスク。

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

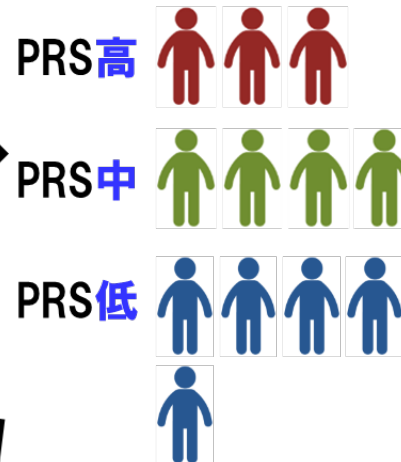


Polygenic Risk
Score推定



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

集団サンプル
層別化モデル



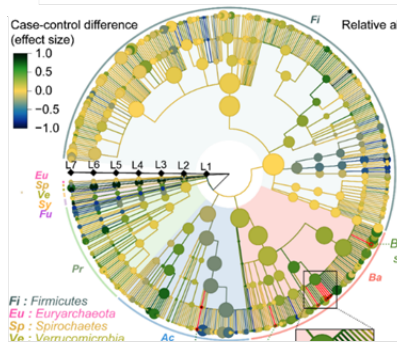
ゲノム
個別化予防



標準的
観察手段

前向きコホート・バイオバンク由来の情報・臨床情報

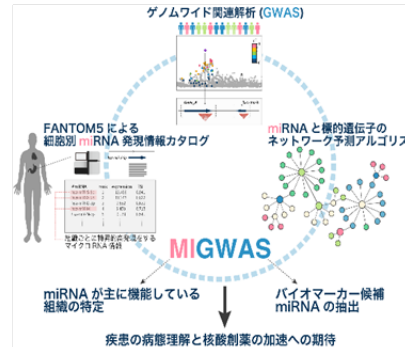
腸内細菌メタゲノム



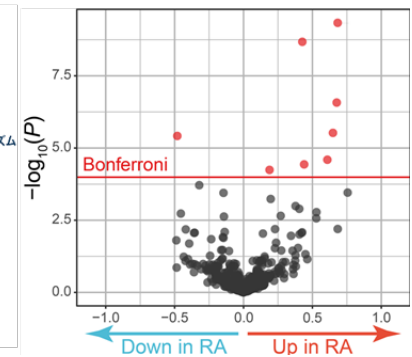
臨床情報・生活習慣

- 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



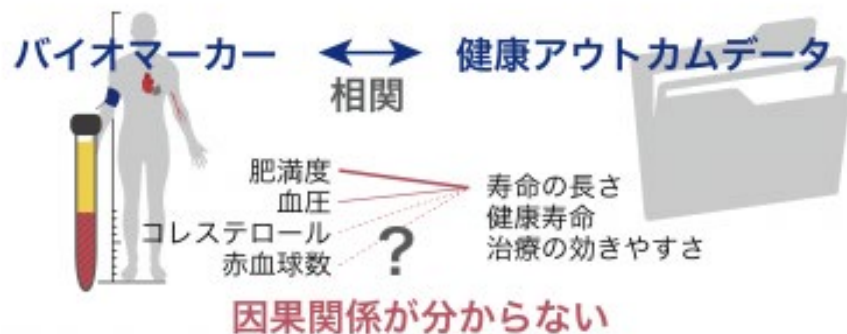
メタボローム



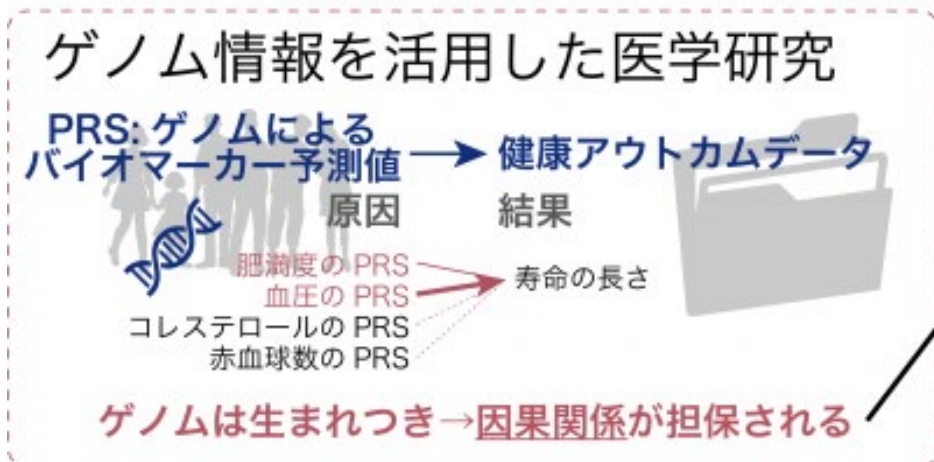
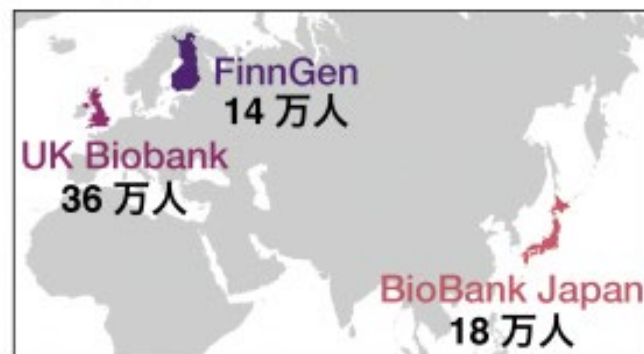
• 加齢に伴い集団中に高頻度で発症する疾患では(例:生活習慣病)、PRSに基づく人生
キャリア早期の高リスク群同定、予防医療、個別化医療が医療経済的にも効果的。⁹

Polygenic risk scoreの健康寿命研究への応用

これまでの医学研究



世界3カ国 70万人のゲノムデータに応用



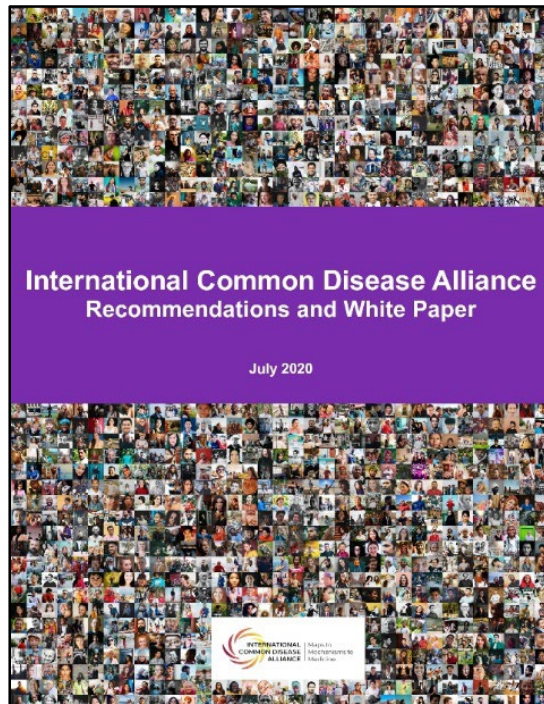
治療により健康アウトカムの改善が期待できる
バイオマーカーを特定

- 国際バイオバンク67万人のPRS解析を実施。健康の指標かつ治療可能なバイオマーカーのPRSと寿命との因果関係を人種横断的に調べる遺伝統計解析手法を開発。
- 高血圧・肥満が健康寿命を短縮し、その治療が寿命延伸に貢献する可能性を示した。
- ゲノム情報の因果性とPRSを結びつける新たな可能性を開拓。

国際バイオバンク連携におけるガイドライン策定

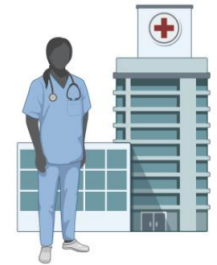
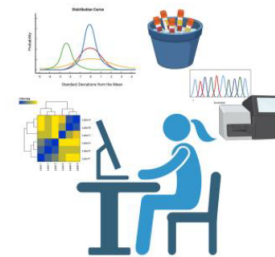
国際ゲノム政策提言組織
ICDA White paperの発表

ICDA PRSタスクフォースを通じたPRS
Best practiceの提言



Future objectives for responsible use of PRSs

Communities + Researchers + Clinicians



Short term (current - 5 years)

Long term (>5 years)

**Establish
Benefits**

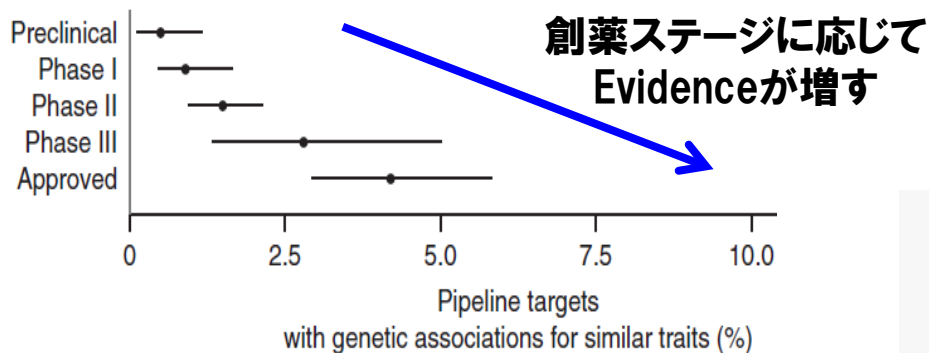
- Determine clinical utility for diagnostic refinement, risk prediction
- Quantify cost effectiveness for specific-use cases and across health systems
- Complement clinical-based lifestyle recommendations

- Adopt standards within professional societies that make PRS risk information actionable
- Create internationally federated informatic platform for PRS implementation to standardize data workflows and clinical pathways

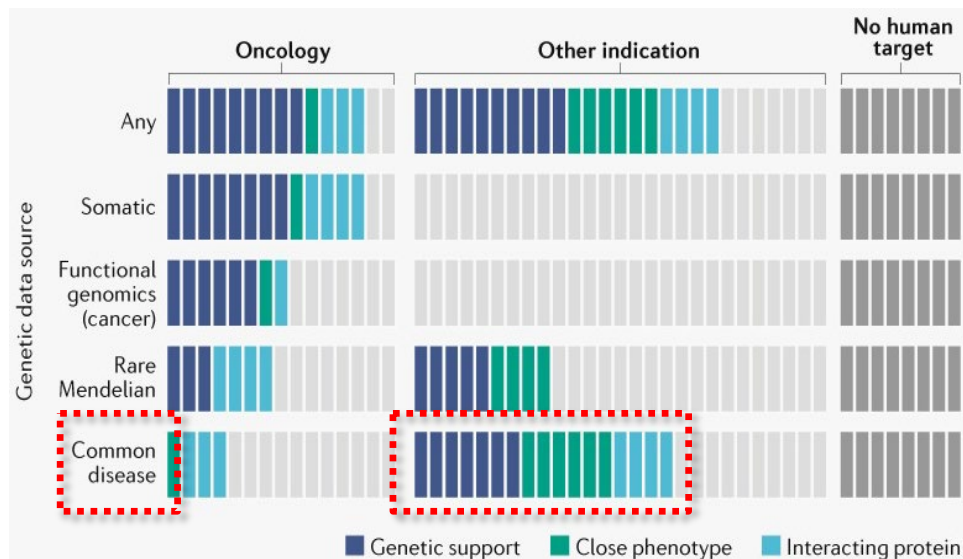
- PRSの社会実装について、国際的にガイドライン策定が進行中。
- 国際ゲノム政策提言組織ICDAのExecutive committeeとして、ヒトゲノム研究の今後を論じたWhite Paper発表や、PRS社会実装に向けたBest Practice提言を行った。

何故、創薬プロセス効率の改善が重要なのか

「Genetic Evidence」



FDA承認治療薬における疾患ゲノム情報の根拠



産官学民によるバイオバンク運営



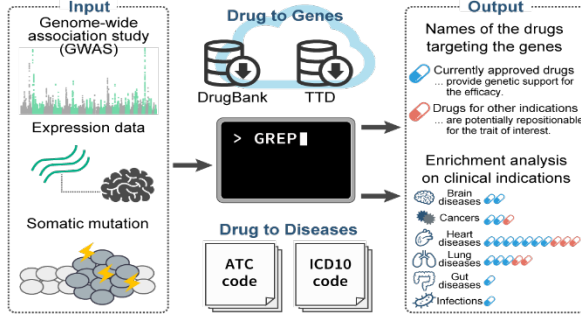
- 疾患ゲノム情報は創薬プロセスを効率化する(例:AstraZeneca社、GlaxoSmithKline社)。
- データアクセスを条件に、製薬企業から出資を募るバイオバンク運営(FinnGen)。
- 2021年に米国食品医薬品局(FDA)に承認された50の治療薬のうち、2/3の治療薬は疾患ゲノム情報の根拠あり。ゲノムワイド関連解析の活用の重要性が示唆された。

疾患ゲノム情報に基づく新規創薬スクリーニング

バイオバンク連携時代のゲノム創薬Best Practiceの提唱

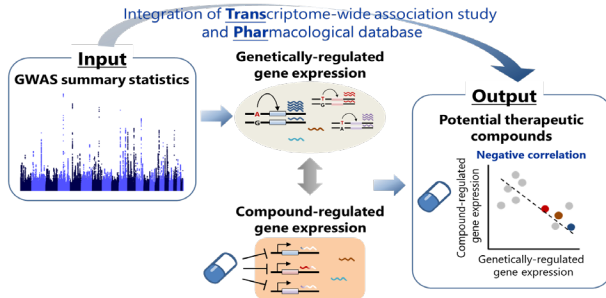
GREP:

Genome for REPositioning drugs



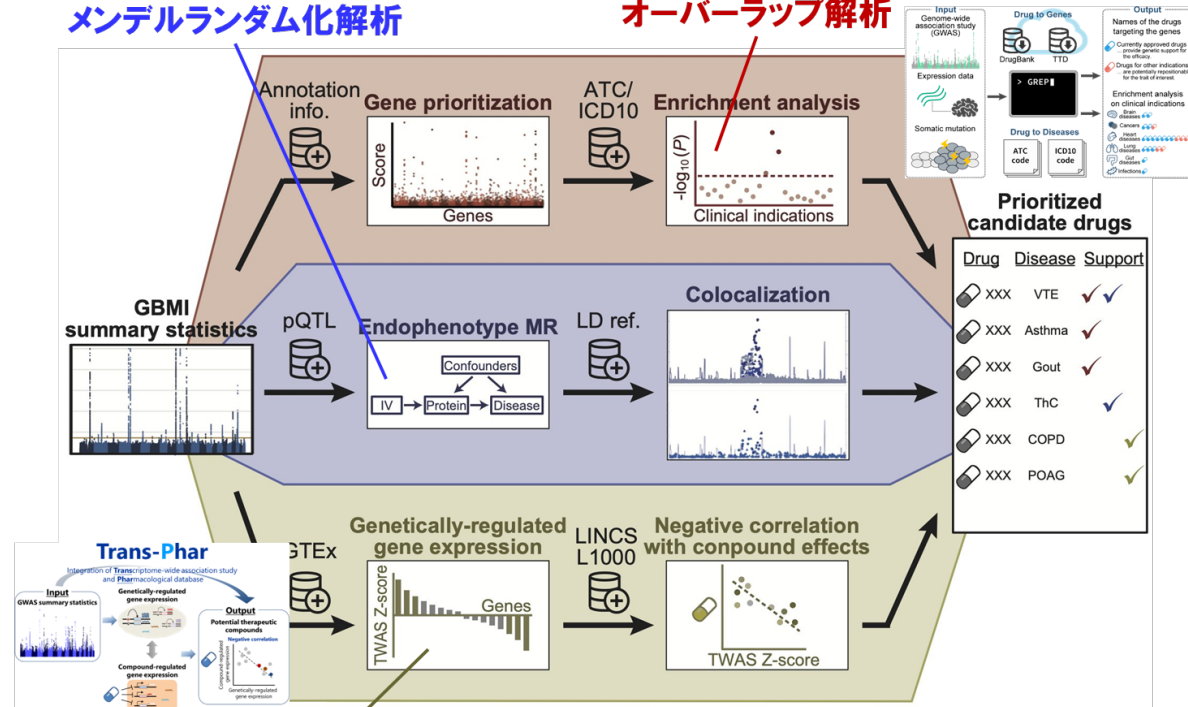
Trans-Phar: Integration of TWAS and Pharmacological databases

Trans-Phar



中間形質QTL解析 & メンデルランダム化解析

疾患感受性遺伝子-創薬標的遺伝子 オーバーラップ解析



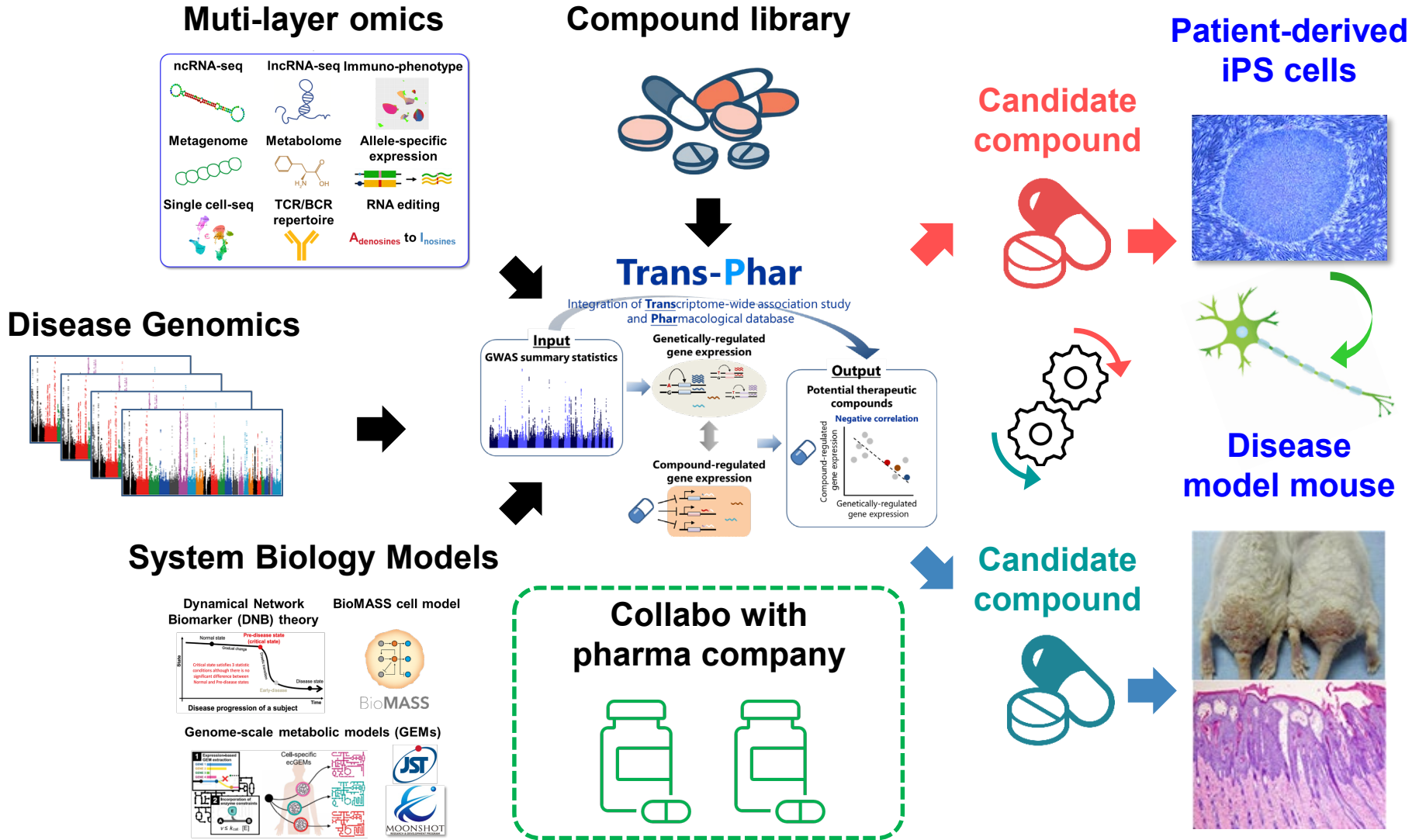
Prioritized candidate drugs

Drug	Disease	Support
XXX	VTE	✓✓
XXX	Asthma	✓
XXX	Gout	✓
XXX	ThC	✓
XXX	COPD	✓
XXX	POAG	✓

TWAS・化合物添加遺伝子発現プロファイル変化

- 疾患感受性遺伝子を直接標的とした治療薬に注目することで、**ドラック・リポジショニング**候補探索の効率化が可能に。自動リポジショニングソフトウェアを開発。
- GREP: WHO ATC分類を活用。Trans-Phar: 遺伝子発現プロファイル変化を考慮。
- バイオバンク連携GWAS時代における**ゲノム創薬のBest Practice**を提唱。

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



- 疾患ゲノミクス解析からin silico創薬標的スクリーニング → in vivoの実装へ。
- 患者由来iPS細胞への添加実験を、患者層別化と共に。モデル生物を用いた実験系。

次代を担う若手研究者の人材育成

遺伝統計学・夏の学校の自主開催
(2022年度はオンライン開催、300名が参加)

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
夏の学校
@大阪大学 ONLINE**
参加費：無料

夏の大阪で、ゲノム解析の世界に触れてみませんか？

日時：令和4年8月26日(金)～8月28日(日)
(原則として、三日間通じて参加をお願いします)

場所：オンライン形式でのWeb開催
(新型コロナウイルス対策の観点より、本年度は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事があります)

時間帯	一日目 (8/26)	二日目 (8/27)	三日目 (8/28)
午前 (9:00-12:00)	遺伝学入門	統計ソフトR入門	ゲノムデータ解析演習
午後 (13:00-16:30)	Linux入門 プログラミング入門	ゲノムデータ解析演習①・②	③・④・⑤・⑥

(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_2022.html
申込み : office@sg.med.osaka-u.ac.jp (担当：岡田 随象)

共催：東京大学大学院医学系研究科遺伝情報学、理化学研究所生命医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、合機ムーンショットプロジェクト
AMEDゲノム医療実現推進プラットフォーム事業GRIFIN



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gzip形式で圧縮
gunzip	gzip形式のファイルを解凍
less	ファイル内容を一画面ごとに表示
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医学部2回生の正規カリキュラムとして開始。
- ・東京大学医学部においても実習の導入を開始。

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。ベンチャー・スタートアップの遅れが、オミクス解析技術における技術開発から現場導入までのタイムラグ延長へ(例:シングルセル解析・Spatial transcriptome)。
- 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・解析基盤の遅れ。
→ 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。

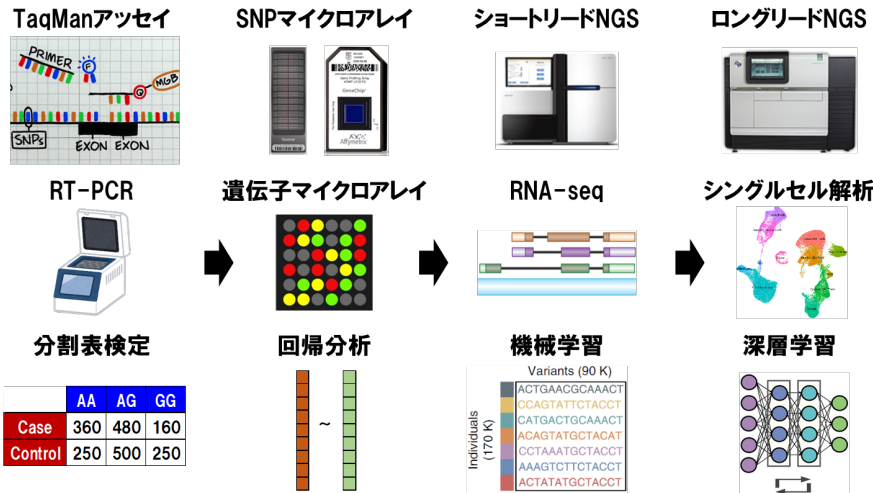
世界に勝つための戦略

- **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」目指す時代に突入。**
 - **学会・大学院の英語化推進による国際人材・国際競争力・大学ランキングの底上げ。**
 - **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。**
 - **数理科学、情報科学、統計学分野の専門家の参入とキャリアパス形成をサポート。**
 - **大規模ゲノムデータ・バイオバンクデータの公開・分譲促進。**
- **データ公開・分譲の遅れが、研究成果・国際イニシアティブの逸失だけでなく、若手人材の参入障壁となっている。**
- 大規模コホート・バイオバンクがデータ公開のリーダーシップをとることが重要。**
- データ所有権の属人化を無くし、組織としての運用体制へ。**
- 元データ公開原則の徹底(GWAS:intensity data/idatファイル、WGS:fastqファイル)。**

遺伝統計学の今後の展望

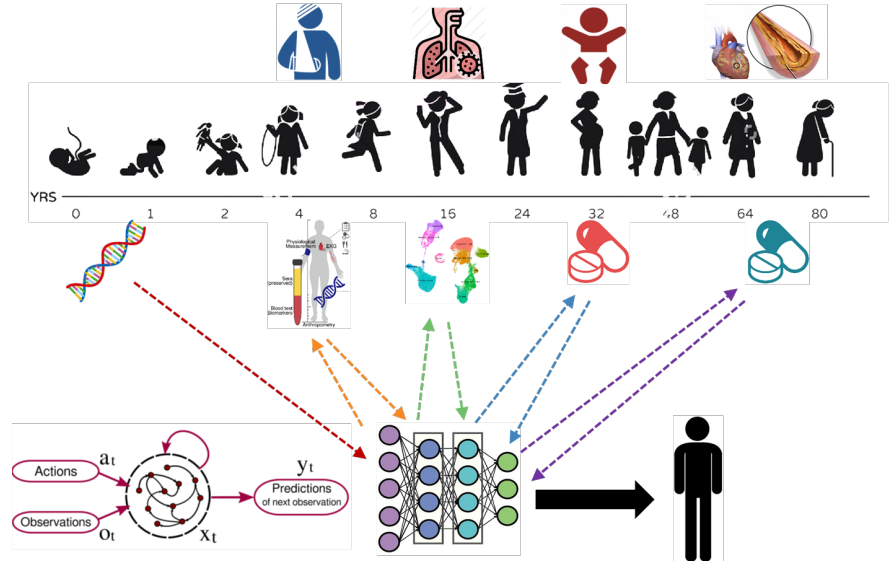
- 誰でも世界最高峰のデータにアクセスできる時代、に即した戦い方が必要。
- 医学、情報学、数理科学、生物学、疫学、多彩なバックグラウンドを持つ専門家チームが、最先端解析技術の導入を通じて、個の技と総合力で勝負していく。
- 「疾病の解明や予測」から、「疾患を制御する」学問への変革。
- データ解析研究のコスト認識に配慮した予算配分。
- 海外資本の柔軟活用（例: Finngen Partners, UK biobank Pharma Proteomics Project）。
- バイオバンク資源の海外研究者への「安全な」開放（例: Genomics England）

Wet・Dry双方の最新解析技術の先進的導入



- 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。

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。