

2023年11月20日

ライフサイエンスの 潮流

鎌谷洋一郎



東京大学大学院
新領域創成科学研究科



賢者は歴史に学ぶ？（私家版GWASの歴史）

研究成果等

手法・理論



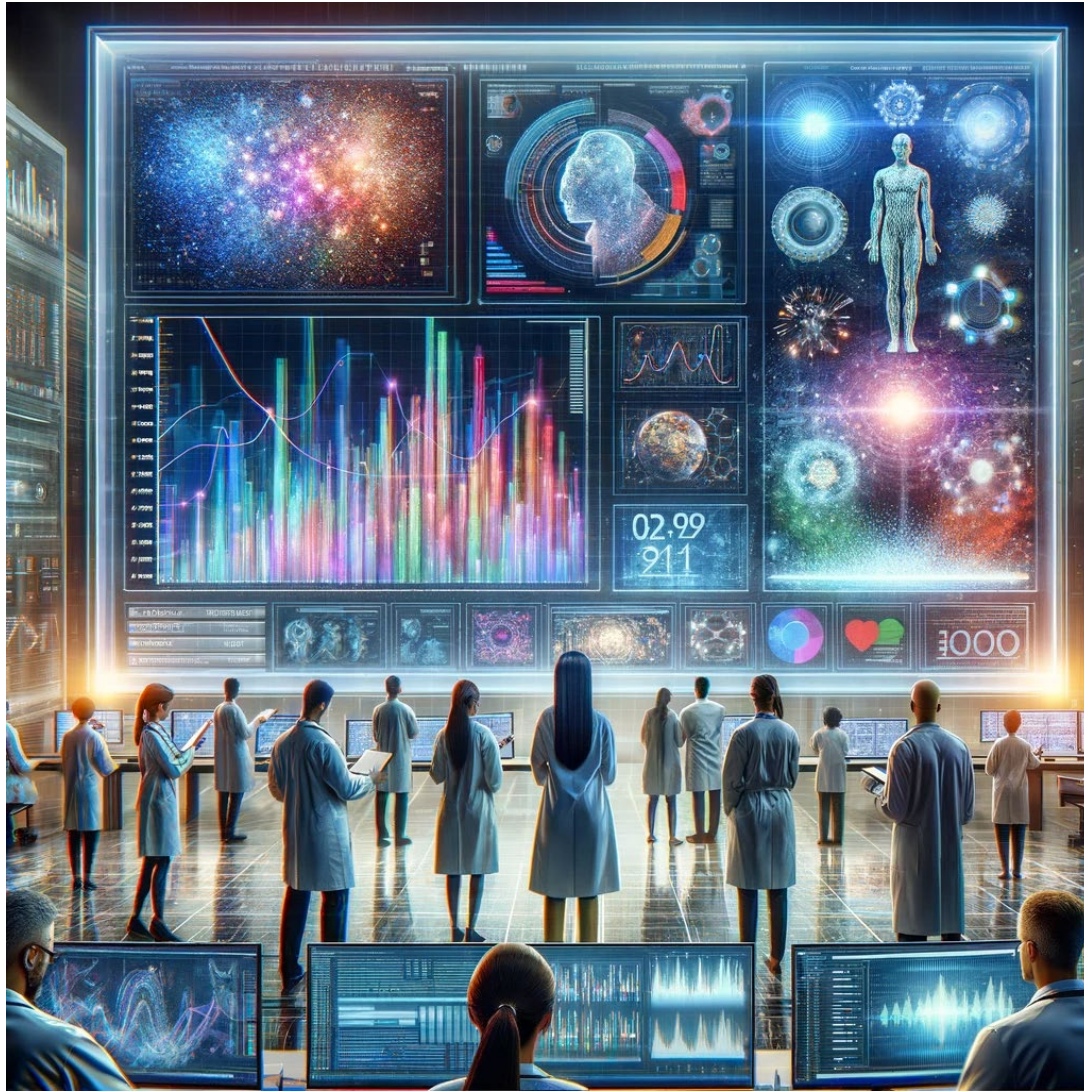
		1919年	FisherがPolygenic modelを提案
		1995年	Genomic Control (GWASのバイアス補正)
1996年	deCODE genetics設立 (アイスランド)	1996年	RischとMerikangasがGWASがよいと予測
2002年	理研が最初のGWASを報告 UK Biobank開始		
2003年	ヒトゲノム解読 BioBank Japan開始	2003年	GWASの検出力 (Chapmanら)
2004年	China Kadoorie Biobank開始	2006年	EIGENSTRAT (PCAで集団構造化を補正)
2007年	WTCCC論文 (GWASの方法論を確立)	2007年	IMPUTE (欠損遺伝型を予測し、50万→数百、数千万バリエーションのゲノムデータに増幅)
2008年	GWAS catalog開始 「見つからない遺伝率」		
2010年	「SNPは身長分散の45%を説明」 GIANTが18万人のGWAS、180座位	2010年	EMMAX、GEMMA (混合線型回帰モデル) MAGENTA (個人情報なしでサマリーデータからパスウェイ解析)
		2011年	GCTA (ゲノムデータから遺伝率を計算) MANTRA (集団横断的メタ解析)
2012年	英国キャメロン首相 “100,000 Genomes Project” ENCODE project論文	2012年	GCTA-COJO (サマリーデータから条件付き解析)
2014年	\$1000ゲノムの達成 FANTOM5		
2015年	米国オバマ大統領 “Precision Medicine Initiative” GTExの一連の論文 Roadmap Epigenomics Project	2015年	LDスコア回帰 (GWASから遺伝率を計算) LDpred (ベイズ型PRSを計算) PrediXcan (GWASとRNA-seqを統合解析)
2017年	SG10K開始 (シンガポール)	2016年	FINEMAP (生物学的原因バリエーションをベイズ推定)
		2017年	MR-Egger法 (メンデルランダム化の一手法) FUMA (GWASの生物学的意義を簡易に提供するインターフェース)
2018年	UK Biobank論文 PRSは単一遺伝子変異と同等の予測能 (Kheraら)	2018年	BOLT-LMM (バイオバンクレベルの高速MLMA) SAIGE (高速に症例対照インバランスを補正)
2020年	gnomADデータベース発表	2021年	REGENIE (高速な機械学習ベースのGWAS)
2021年	国際バイオバンク連携 (日本からは岡田らが参加)		
2022年	GIANTが540万人のGWAS、1万座位		
2023年	UKB-PPP (プロテオームGWAS) UKB 50万人WGS公開予定		

何がわかったか

- 大サンプルのGWASは1疾患・形質あたり **多数の遺伝因子** を発見できる
- 大サンプルのGWAS結果は、**過去に知られている医学的・生物学的知見を反映** しており、さらに未知の新規の情報も提供している
- GWASを元に **疾患発症を予測** でき、手法は成熟している。ただし、民族を超えて予測する際には懸案事項がある。
- 形質によるが、WGSによってわかる **レアバリエント** が、SNPによって見つからない遺伝率Missing heritabilityを大きく説明していることがある
- メンデルランダム化手法が進歩し、GWASを利用してさまざまな **因果推論** を頑健に実施可能となった。
 - 例えばUKB-PPP結果は、「そのタンパク質レベルを下げると病気になりやすくなるか？なりにくくなるか？」を頑健に推定できる→創薬ターゲット。
- GWASからの **Drug repurposing** は有望
 - 世界初のCRISPR治療薬VERVE101のターゲットだって、遺伝的関連のある遺伝子だ
 - *BCL11A* もCRISPR治療薬のターゲットだが、これはGWASが検出した遺伝子
- PRSの有効性を検証する **臨床試験** が複数始まっている

* Visscher et al. Five years of GWAS discovery. AJHG 2012. Ten years of GWAS discovery. AJHG 2017. 15 years of GWAS discovery. AJHG 2023. を参考

バイオバンクの世界的な広がり



DALLE3

- バイオバンクは世界中でどんどん広がっており、広がりを続けている
- ヒトからデジタルデータを取得し多数の人から得ることで、ビッグデータとして解析する
- イギリスの流れ
 - WTCCC (SNP 17,000人)
 - 100,000 Genomes Project (WGS 10万人)
 - UK Biobank (WGS 50万人)
 - Our Future Health (SNP 500万人)
- アメリカ
 - All of Us (WGS 100万人)
 - MVP (SNP 100万人)
 - TOPMed (WGS 18万人)
 - RGC Exome (WES これまで200万、+100万)
- シンガポール：PRECISE (WGS 10万人)
- カタール：Qatar Biobank (10万人目標)
- 南アフリカ：WGS 11万人目標の精密医療

ゲノム・オミックスの未来（まとめ）

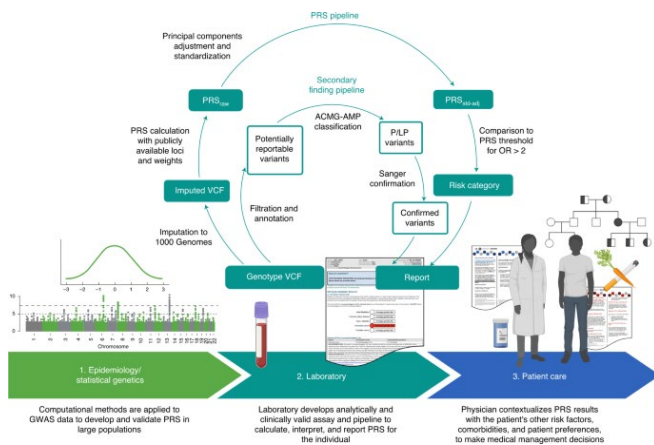
- ヒトをデジタル化したビッグデータをさらに大きく取得、統計・AI解析していく。ゲノムだけでなく、エピゲノム、トランスクリプトーム、プロテオーム、メタボローム、グライコーム、メタゲノム。シングルセル解析。新しいことを解明するために新しいデータを取得する。まだまだ出てくるだろう。患者データについては病院情報（EMR）と統合。これについて
 - どんどん計測していく。新たな計測技術を開発する。
 - どんどん解析していく。新たな解析手法を開発する。
 - これらを医学者・生物学者・生物情報学者のみならず、数学、物理学、工学、情報学の研究者グループと共同で。
 - 医療の視点を忘れずに。
 - また、心突然死の人、薬剤副作用の人、など本人やご家族に負担なしでデータを将来のために取得する方法を作るべき
- 最終的な目標は、コンピュータ空間におけるデジタルヒトデータ再構築と思います
- これを進めるため、アメリカはオバマ大統領、イギリスはキャメロン首相が主導した。「研究者」レベルにあまり大きなお金と決定権がついて良いことはないように感じている（もしくは、そこに関わっている間は研究から離れるべきか）。若手の研究者の一人として、政治・省庁の主導を求めたい。
- 希少遺伝性疾患やがんのシーケンスは診療段階、ありふれた疾患のポリジェニック・スコアは診療に近い可能性。患者返却しながらデータを研究活用する。
 - 返却に合わせ、医療従事者や患者がビッグデータを理解する必要がある
 - 上記の他、法律、倫理、そして教育の専門家とも連携していく必要がある
- このようなヒト生体ビッグデータ推論結果は、実験研究の仮説立案や結果の補強に活かすことでライフサイエンス研究を加速する。実験の重要性が失われるところは少しもなく、むしろ力を合わせていくことが重要。

短期的には何をすべき？

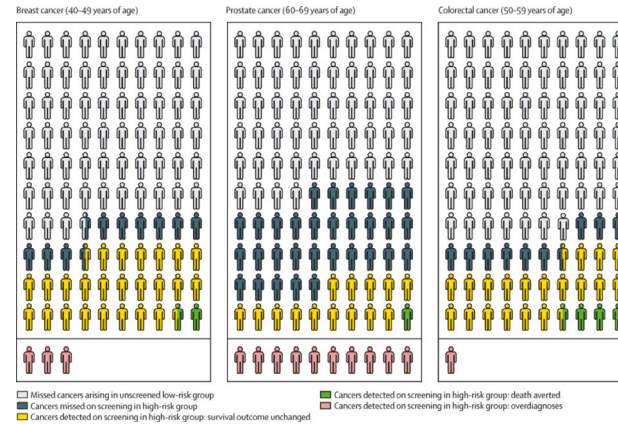
- 複数の選択肢がある
 - ヒト検体ビッグデータ解析をやめる（米中に敵わないから）
 - **SNPアレイ**でどんどん人数を増やす
 - おそらく間違いなく、身長以外のほぼ全ての形質で、欧米人や中国人に加えて日本人をやることで発見される新規の知見はまだまだ出てくる
 - 作業的で、科学研究っぽくないかも
 - 診療応用を同時に考えるなら、単一遺伝因子を取れないのが難点（**SNP+WES**という戦略をとっている米国の診療応用グループはある）
 - **WGS**でどんどん人数を増やす
 - これまでまだあまりわかっていないレアバリエーションをもっと解明できる。
 - **CHIP**、テロミア長などの**SNP**以外のゲノム因子を幅広く解析できる
 - 高価で、**SNP**ほどの検出力を出せない可能性がある
 - 診療応用を同時に考えるならこちらの方がベター（単一遺伝因子を網羅できる）
- オミックスを**広く**取る
 - つまり大人数
- オミックスを**狭く、深く**取る
 - つまり、人数としては比較的少ないがより多層に

中期的目標：診療と連動するデータ収集

- がん・希少難病のための診療応用：すでにやっている
- その他の疾患（ありふれた）：これらも同様に、将来的にはPRSの診療実装ができるはず。病院で幅広くゲノムデータ取得を支援し、予防や診療実装を目指した臨床試験の実施への支援。ただし、がん・難病よりも研究側面が強いことは明らか（患者さんへの返却メリットがそれほど大きくはない）。ただし、なくはないので



GenoVA研究：PRSの診療実装可能性の評価。早期診断、層別化治療等々を評価
Nat Med 2022



シミュレーションによると、PRS層別化検診により英国で乳癌、大腸癌、前立腺癌による年間の癌死を102人、188人、158人減らせると試算
Huntley et al. Lancet Oncology 2023

- そしてAI診療へのPRS組み込みを検討。
- 返却に合わせ、医療従事者や患者がビッグデータを理解するための教育
- 上記の他、法律、倫理、そして教育の専門家とも連携
- この際に取得するEMR、ゲノムデータを研究利活用
- 血液や組織検体を保管し、研究者が使用したいときに提供できるようにする（バイオバンク）

長期的には

- 100年先を見通す研究者を輩出するにはどうすれば良いか？
 - 横断的に活躍できる天才を支援するべき
 - フィッシャーは数学者・遺伝学者
 - 鎌状赤血球で分子病の概念を提唱したポーリングは、博士号は数理解物理学
 - ワトソンとクリックのクリックは物理学者
 - ヒトゲノム論文筆頭のエリック・ランダー氏は経済学者→遺伝学者
- 今よりその次に来るものは何か？
 - （短期的の方かもしれませんが）データインテグレーション。核とリボゾーム（ゲノム+エピゲノム→トランスクリプトーム→プロテオーム）だけにおわらない。シグナリング、生化学的代謝経路、全身の免疫・炎症、脳などをデータ統合していく。。

国際展開について

- 基本的に、オープンサイエンスを実施する国が賞賛されている
- それぞれの国ごとの戦略があり、世界の中での日本の位置どりとしてはどう考えるか。
- 私はオープンサイエンスに舵を切るのが、日本のために良さそうかと思いました。