

NCBN NEWSLETTER

National Center
Biobank Network

ナショナルセンター・
バイオバンクネットワーク

2024.01.25 Volume 9 No. 2

Feature

NCBNバイオバンクの横断的利活用に向けて

日本医療研究開発機構 (AMED) の「ゲノム医療実現バイオバンク利活用プログラム」の課題の1つである「ゲノム医療実現推進のためのバイオバンク・ネットワーク構築とバイオバンク利活用促進に関する研究開発」では、全国のバイオバンクを横断する試料・情報の利活用促進環境の整備が進められており、参加機関からの試料・情報提供について、ワンストップの窓口をつくることを目指しています。

NCBNは発足当初から、横断検索システムや統一された申請窓口を実装してきており、この研究のよいモデルとなっています。一方で、6つのNC (国立高度専門医療研究センター) のバイオバンクを横断的に利用するには、申請書や手続きなどの違いがあるため、これらの違いを統一する上での実務的課題を明らかにし、解決を図る必要があると認識しています。

そうした中で、6NCの試料・情報を利用する研究プロジェクトが2020年にスタートしました。それは、国立高度専門医療研究センター医療研究連携推進本部 (JH) が支援する「ライフコースのメンタルヘルス課題についての疾患レジストリを活用した、6NCによる共同研究プロジェクト」でした。この研究プロジェクトは、患者さんの試料 (血漿・血清) と情報 (身体疾患に関する臨床情報とメンタルヘルス評価) をレジストリに登録し、審査を経て解析し、試料中の分子と情報との関連を検討して、心身相関のメカニズムの解明及び疾患横断的な関連マーカーの開発を目指しました。この研究で6つのNCバイオバンクの試料・情報を横

NCBN中央バイオバンク事務局 野入英世
国立循環器病研究センター (NCVC) 富田努
国立精神・神経医療研究センター (NCNP) 服部功太郎

断的に利用する際の課題を解決するための、バイオバンク実務者による会議が重ねられました。

その成果として、6NCそれぞれからの試料・情報の同研究プロジェクトへの提供を1通の共通申請書で実施することができました。一方、研究計画立案時からバイオバンクの関係者が関わって調整することが極めて重要であることが改めてわかったため、以降のプロジェクトに活かせるように、NCBNが基盤となって研究者と調整することで試料・情報の利用が円滑に進むような仕組みを整備しました。同時に、各NCの手続きや倫理委員会の判断の違いも、これらの過程で再確認できました。

そこで、このほどNCBNは外部委託による第三者機関のヒアリングを経て、6NCバイオバンクの横断的利活用に向けた課題と今後のあり方を報告書に取りまとめました。その中で同研究プロジェクトの経験を整理すると同時にNC間の類似点と相違点を可視化したほか、ユーザーへのインタビューによりNCBNへの期待と課題を整理すると共に、NCBN以外の国内外のバイオバンクについても同様の目線で調査しています。

それらを踏まえた今回の報告書には、今後の方向として、SMO (治験施設支援機関) の設置と、利用者に対するe-learningの提案が含まれています。NCBNは、報告書の内容を鑑み、6NCのネットワークを活かしたさらなる利活用の推進やバイオバンク・ネットワークにおける展開に向けて様々な取り組みを検討・実施していく予定です。

河合洋介先生に聞く

コントロール群全ゲノム解析から見えてきたもの



河合洋介先生
国立国際医療研究センター
ゲノム医科学プロジェクト
副プロジェクト長

NCBNのゲノム解析に協力

— NCBNは2020年度に、国が進める「全ゲノム解析等実行計画」に協力してコントロール群9,850人の全ゲノムシーケンス(WGS)解析を行いました(NCBNニュースレターVol. 8 No. 1参照)。先生はこの解析にどのように関わられたのでしょうか？

河合：このWGS解析は、NCBNとしてNCが協力して行う初めてのゲノム解析でしたので、参加した5つのNC間でばらつきのない、高品質の解析データが得られるようにするため、中央バイオバンク事務局に協力して共通のプロトコル(手法)を作成しました。各NCから品質を保ったまま試料(検体)を解析業者に渡し、解析データを中央バイオバンクが受け取って各NCに渡すまでのフローをつくるのには苦労しました。具体的には、検体の管理・追跡システムをつかって、各検体がどの段階にあるか、解析がどこまで進んでいるのかを把握できるようにしました。またWGS解析データはサイズが大きいため、データの共有やバックアップのためのリソースも確保しました。

— そのようにして得られたWGS解析データを、NCBNから提供できるようになったのですね。

河合：はい。このデータはがんと難病のコントロール群として利用される予定ですが、得られたデータを早く有効に活用していただくため、様々な研究機関との共同研究という形で提供が始まっており、合計で20件(2023年12月現在)にのぼっています。

— さらに、先生はコントロール群のWGS解析データを用いたご研究を主導され、2023年12月に論文発表*1されました。どのような研究をされたのでしょうか？

河合：今後の利活用を見据えて解析で得られたデータの品質評価を行いました。また、5つのNCが東京都、愛知県、大阪府にあり、解析対象となった登録者の出身地が全国に散らばっているという特徴を活かし、今回のデータから日本人の集団遺伝学的特性*2を明らかにする研究を行いました。

高品質でNC間のばらつきの少ないデータが得られた

— データの品質評価はどのように行われたのでしょうか？

河合：まず、5つのNC間で解析品質にばらつきがあるかどうかを評価しました。WGSでは、ゲノムを数百塩基の断片にして配列を読み取り、個々の断片を参照ゲノム*3に照らし合わせてその位置を決め、全断片の情報を合わせることで全体の塩基配列を得ます。解析品質の評価では、平均深度(断片を集めたものが全ゲノムを30回以上カバーしているか)、マッピング率(各断片が参照ゲノムにちゃんとあてはまるか)、平均挿入サイズ(断片サイズがある程度揃っているか)が重要です。この3つの値を検体ごとに求めたところ、5NCとも品質が高く、NC間でのばらつきはほとんどありませんでした(図1)。検体が唾液の場合、少し変動がありました。これは口腔内の雑菌など様々な由来のDNAが混じったためと考えられ、以後の解析への影響はほとんどありませんでした。

*1: Y. Kawai et al., Exploring the genetic diversity of the Japanese population: Insights from a large-scale whole genome sequencing analysis, PLOS Genetics 2023. DOI:10.1371/journal.pgen.1010625

*2: 集団遺伝学とは、ある集団の遺伝情報を統計的に解析することで、その集団の遺伝的な特徴の時間的変化や、他の集団との違いを明らかにする学問分野。

*3: 全ゲノムの塩基配列を解析するときに各断片の位置を決めるための基準となる配列。2003年に初めて決定されたヒトゲノム配列をもとにしている。

また、解析データを、同じ機種種のシーケンサーを使った国際1000人ゲノム計画*4のデータと比較したり、いくつかの方法でエラー率を推定したりした結果、エラー率はかなり低いことがわかりました。

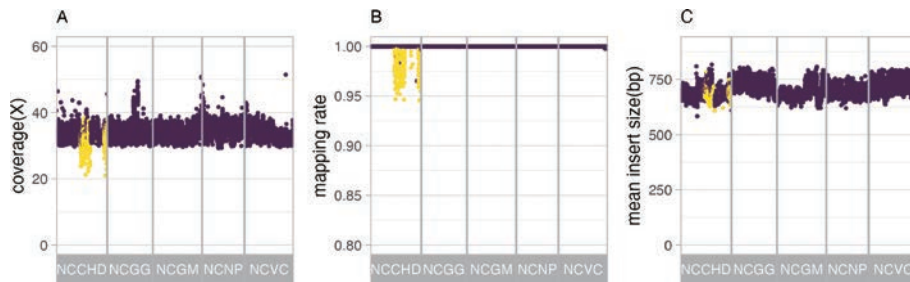


図1: 5NCの解析品質の評価。A(平均深度)、B(マッピング率)、C(平均挿入サイズ)とも品質が高く、ばらつきはほとんどなかった。黄色で示したのは唾液由来のDNA。*1の論文より。

日本人の集団遺伝学的な特徴も明らかに

——集団遺伝学のご研究ではどのような解析を行い、どのようなことがわかったのでしょうか？

河合: 今回得られたデータと、国際1000人ゲノム計画の2,504人のデータと統合して解析し、どのような集団構造が見られるかを調べました。主成分分析*5により一人を一つの点としてプロットしたところ、NCBNの集団は独自のクラスターを形成しており、国際1000人ゲノムのうちの東アジアの大陸集団(漢民族、キン族、ダイ族)とは遺伝的に分化していることや、日本人集団内の遺伝的多様性を確認しました(図2)。これは過去の研究でもいわれていたことですが、今回の解析でより明確になりました。

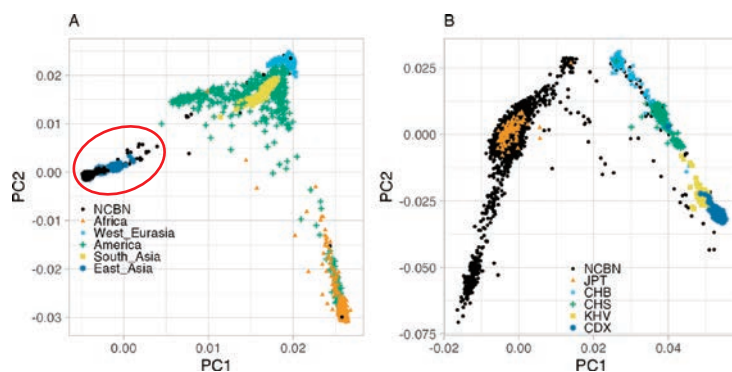


図2: WGS解析データの解析で明らかになった集団構造。AはNCBNのデータと国際1000人ゲノムのデータを合わせて解析した結果。NCBNと東アジアが一つの集団(赤丸)をなしている。Bは、赤丸の集団をさらに解析した結果。NCBNとJPT(国際1000人ゲノムのうち東京の日本人)からなる集団と、CHB(北京の漢民族)、CHS(中国南部の漢民族)、KHV(ベトナムホーチミンのキン族)、CDX(中国シーサンパンナのダイ族)からなる集団が分化している。*1の論文の図を改変。

また、ゲノム情報からは、過去の人口が世代ごとにどう変化してきたかを推定することもできます。これにより、本土と沖縄では、集団の過去の人口動態が異なることを明らかにしました。

さらに、過去の研究で、東アジア人に固有な遺伝子で正の自然淘汰を受けていると報告されている遺伝子にも注目しました。中でも、アルコール代謝に関連する遺伝子(ALDH2, ADH1B)、メラニン合成に関連する遺伝子(OCA2)、脂肪酸代謝に関連する遺伝子(FADS)の集団内での遺伝子頻度が約3万年前から現在までにどう変化してきたかを推定したところ、これらの遺伝子は2万5000年から1万年前に集団内での遺伝子頻度が急激に増えていました。つまり、正の自然淘汰を受けているという先行研究の結果が支持されたのです。

——集団遺伝学的な解析ではずいぶんいろいろなことがわかるのですね。

河合: これらに加えて、稀なバリエーションの影響の評価やヒト白血球型抗原遺伝子(HLA)の解析も行いました。今回のWGS解析で得られたデータは全ゲノム解析等実行計画の中で活用されることはもちろんですが、日本国内の地域的多様性を考慮した集団遺伝学的解析、疾患解析にも役立つと期待しています。

*4: 世界の主要な集団を対象とした全ゲノム解析を行う国際共同研究。International 1000 Genomes Project.

*5: 変数を減らすことでデータ分布の特徴を抽出しやすくする統計解析方法。

今年度の学会出展

2023年5月に新型コロナ対策の行動制限が撤廃され、海外も含め多くの学会に出展することができました。今後も対面での広報活動を充実させ、利活用の拡大につなげていきます。

学会名	開催地	開催日
済 第67回 日本リウマチ学会	福岡	2023/4/24-26
済 第8回 クリニカルバイオバンク学会シンポジウム(日本遺伝子診療学会大会と合同開催)	千葉	2023/7/28-29
済 第55回 日本医療検査科学会	横浜	2023/10/6-8
済 第68回 日本人類遺伝学会	東京	2023/10/11-14
済 American Society of Human Genetics	Washington, DC	2023/11/1-5
済 第70回 日本臨床検査医学会	長崎	2023/11/16-19
済 第46回 日本分子生物学会	神戸	2023/12/6-8
済 第97回 日本薬理学会年会・第44回 日本臨床薬理学会学術総会(同時開催)	神戸	2023/12/14-16
第88回 日本循環器学会学術集会	神戸	2024/3/8-10

コントロール群全ゲノム解析の成果を論文発表

2023年12月7日(米国東部時間)、NCBNの5つのNCが協力して行った全ゲノム解析の結果と、そのデータを利用した研究成果がPLOS Genetics誌に論文発表され、プレスリリースも行いました(詳細は今号の河合先生インタビュー記事参照)。全ゲノム解析に基づく医学研究が世界的に進展する中、NCBNの活動意義をアピールする論文となりました。

<https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1010625>

Catalogue database

NCBNカタログデータベース試料登録情報(2024年1月1日時点)

NCBNの活動にご理解、ご賛同いただきましてありがとうございます。患者さまのご協力により、試料登録数の合計は472,413件(2023年6月12日時点)から490,783件(2024年1月1日時点)へと着実に増えています。生体試料種別、ICD-10コード別の登録数は表の通りです。

生体試料種別登録数一覧

生体試料	GNC	NCC	NCVC	NCNP	NCGM	NCCHD	NCGG
生細胞	28,335	0	28,335	0	0	0	0
血液・血清・体液等	219,418	54,135	62,511	25,021	43,202	1,192	33,357
組織	32,196	18,564	535	10,853	0	1,607	637
核酸	210,619	108,270	52,629	21,362	12,155	1,648	14,555
その他	215	0	0	0	215	0	0
合計登録検体数	490,783	180,969	144,010	57,236	55,572	4,447	48,549

カタログデータベースでは、病名、生体試料種別、年齢・性別などから登録試料を検索することができます。最新の統計もご覧いただけます。

■情報の確認・検索はこちら

PC▶

<http://www.ncbiobank.org/Index>

スマホ▶



ICD-10コード別疾患登録数一覧

ICD10分類	GNC	NCC	NCVC	NCNP	NCGM	NCCHD	NCGG
A00-B99 感染症および寄生虫症	6,102	23	1,402	102	3,692	5	878
C00-D48 新生物	69,715	60,618	3,789	302	3,535	109	1,362
D50-D89 血液および造血系の疾患並びに免疫機構の障害	2,222	11	1,645	45	244	20	257
E00-E90 内分泌、栄養および代謝疾患	21,750	14	16,463	1,033	3,107	15	1,118
F00-F99 精神および行動の障害	11,575	1	1,458	4,849	315	4	4,948
G00-G99 神経系の疾患	26,735	2	6,040	15,279	518	7	4,889
H00-H59 眼および付属器の疾患	5,670	0	2,142	168	1,384	4	1,972
H60-H95 耳および乳突突起の疾患	1,672	0	730	53	106	1	782
I00-I99 循環器系の疾患	31,960	14	24,811	1,325	3,533	8	2,269
J00-J99 呼吸器系の疾患	5,836	64	3,421	97	1,016	10	1,228
K00-K93 消化器系の疾患	12,472	175	6,709	138	4,025	18	1,407
L00-L99 皮膚および皮下組織の疾患	2,701	11	738	27	307	4	1,614
M00-M99 筋骨格系および結合組織の疾患	8,419	42	3,380	428	1,148	21	3,400
N00-N99 泌尿生殖器系の疾患	9,181	66	5,756	147	1,731	8	1,473
O00-O99 妊娠、分娩および産後<前>	1,178	0	667	1	85	424	1
P00-P96 周産期に発生した疾患	319	0	40	8	0	271	0
Q00-Q99 先天奇形、変形および染色体異常	3,321	15	2,646	314	79	233	34
R00-R99 症状、徴候および異常臨床所見・異常検査所見で他に分類されないもの	6,562	14	3,916	289	335	7	2,001
S00-T98 損傷、中毒およびその他の外因の影響	7,921	17	2,887	2,629	515	10	1,863
U00-U99 特殊目的用コード	335	2	71	73	1	1	187
V00-Y98 傷死および死亡の外因	97	0	75	3	1	4	14
Z00-Z99 健康状態に影響をおよぼす要因および保健サービスの利用	15,961	10	8,651	1,829	711	882	3,878

ナショナルセンター・バイオバンクネットワーク(NCBN)は、6つの国立高度専門医療研究センター(NC)が「新たな医の創造」に向けて個々の疾患専門性を尊重しつつ、ネットワーク型・連邦型の組織形態で運営するバイオバンク事業です。

NCBN中央バイオバンク事務局

〒162-8655

東京都新宿区戸山1-21-1

国立国際医療研究センター内

Tel:03-5273-6891

mail:secretariat@ncbiobank.org

<https://www.ncbiobank.org/>



National Center
Biobank Network