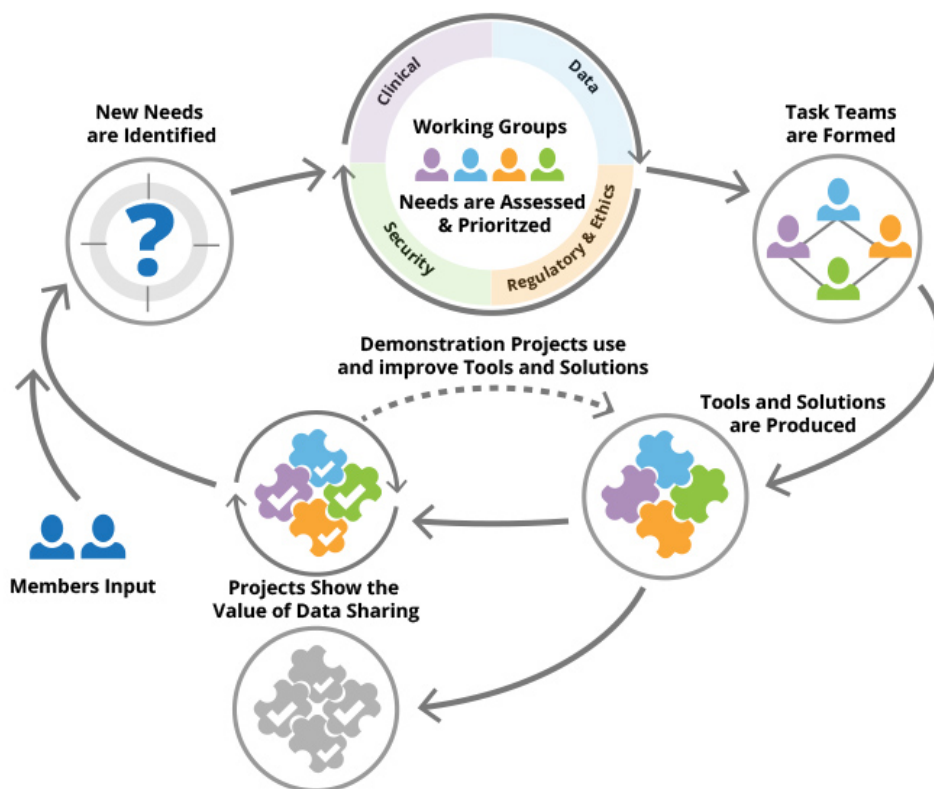


Global Alliance for Genomics and Health (GA4GH)の活動について

国立研究開発法人科学技術振興機構
 バイオサイエンスデータベースセンター

2015年6月9-11日にオランダのライデンで the 3rd Plenary Meeting of GA4GH が開催された。GA4GH は、『様々な機関が集結し、参加者の自主性を尊重しながらも、ゲノムデータと臨床情報を最大限に統合・活用し、ヒトの健康増進を図る』という Missionのもと、2013年1月28日に開催されたキックオフミーティングを皮切りにグローバルな健康情報や診療情報、遺伝学的情報といったデータの共有を目指した Framework や Best practice の作成および提案を発信している団体である。Data working Group (DWG), Security WG (SWG), Clinical WG (CWG), Regulatory and Ethics WG (REWG)の4つのWGに分かれており、それぞれのWGの中に複数のTask Teamを置き、データ共有に必要な事柄についてTV会議を中心に検討を進めている。同時に、WG間の連携を図ることで、1つのWGでは深く検討できない案件についても、多様な専門家の視点を入れて検討している。



<http://genomicsandhealth.org/about-the-global-alliance/how-we-work>

WG the 3rd Plenary Meetingにおいて検討されていた内容の内、特に Data Access Committee に関与する部分について報告する。

動画とスライドは下記 URL から公開されている：

<http://genomicsandhealth.org/newsevents/events/june10thmeetingpresentations>

1 : CWG

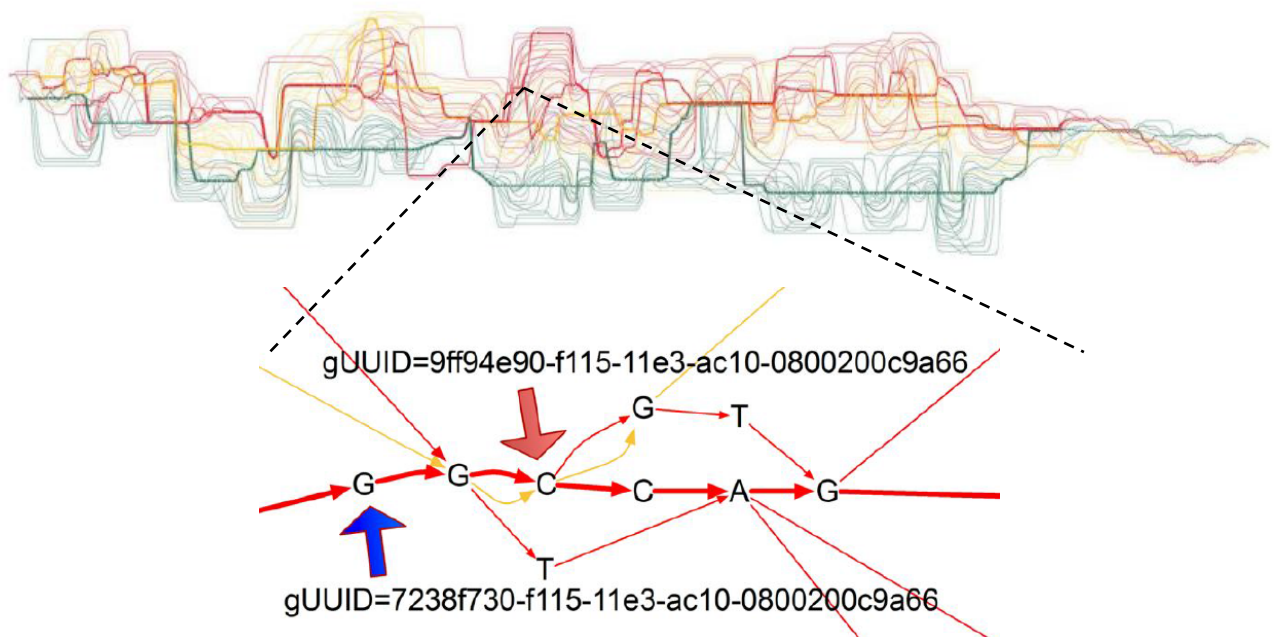
Phenotype データの統一フォーマットの開発（オントロジー）やゲノムデータとのリンク付け方法の確立を目的としている。まずは希少疾患とガンから取り組む。将来的には多因子疾患や感染症にも広げていく予定。

- ・エビデンスに基づいて Genotype と Phenotype の関係性をどのように記述していくかの検討。
（遺伝子変異と薬効・副作用、治療反応性、といった、変異と機能のアノテーション付け）
- ・ガンについては、治療方針決定や試験をできるほど十分な患者を集められないことから、データを共有することは大事。
- ・希少疾患については、信頼できるエビデンスのためにも、国際的にデータを共有する必要がある。
- ・最終的には、医師の意思決定のサポートを目指す。
- ・ゲノムからの個人特定性についての議論もしている。

2 : DWG

データ形式、クラウド環境における安全な保管、ゲノム情報を共有するためのアプリケーション・プログラミング・インターフェース (API) の開発、データを使いやすくするためのアプリケーション開発といった、技術開発を実施。

- ・Reference Graph の開発（Raw data は容量が大きいため、グラフの形で保管・共有できないか？）



<http://genomicsandhealth.org/files/public/6BeaconHaussslerGA4GHLeiden.pdf>

- ・個人情報をクラウドで共有することが大前提であり、最終目的でもある（SWG と連携）。

3 : SWG

データセキュリティ、アクセス制御、監査機能、プライバシー保護について検討している。

- ・ Privacy and Security Policy を作成＞具体的かつ実用的なガイドライン作成を目指す。
- ・ ユースケースの調査、Best Practice の選定を中心に検討していく。
- ・ 秘匿計算の開発
- ・ ユーザー登録についての検討（何の情報をもって利用を許可すればよいのか、自己申告をどこまで信用すべきか）

- ・ クラウドのプライバシーとデータ保護に関する世界初の国際規約：IS027018（2014/8/1）

http://www.iso.org/iso/catalogue_detail?csnumber=61498

- ・ データの分散管理、集中管理について検討。

・ セキュリティ事故が起きた時の対応について検討（定義、検知方法、組織・国の違いを超えてどこまで追及できるか、アタックを受けた時の対応など）。

- ・ 個人同定を禁止するポリシーを、技術的に担保できないか（例：User Tracking）

- ・ iPhone アプリを使ったデータ収集
- ・ 既存ツールの脆弱性テスト

4 : REWG

国際ガイドラインや倫理的な枠組みを作成し、ゲノムデータ・臨床情報の信頼のおける共有を世界規模で活性化させることを目的としている。

・ ゲノムデータ・医療情報を共有するための国際規約の作成（Framework for Responsible Sharing of Genomic and Health Related Data）

<http://genomicsandhealth.org/about-the-global-alliance/key-documents/framework-responsible-sharing-genomic-and-health-related-dat>

・ Consent Policy を作成：これまでは、“共有”がゴールだったが、今後は、どのように Health 関連の情報を研究参加者に返していくための Consent や Access 権についても盛り込んでいく必要がある。

<http://genomicsandhealth.org/files/public/Consent%20Policy%20%28Final%20%2027%20May%202015%29.pdf>

- ・ International data sharing の最大の問題点：他国にデータを出したくない。

・ EU データ保護指令（2016 年に通過し、2018 年ごろから効力を持つようです）や個人情報保護法の改定（それぞれの国の法律）についても意識していかなければならない。

- ・ OECD や UNESCO の国際 Bioethics Committee との関わり方

・ 国際的な多施設共同研究を進めるにあたり、各国の Research Ethics Committee の相互理解を進めたい。研究者パスポート（accreditation）を導入すべきなのか？

- ・ インフォームドコンセントについて

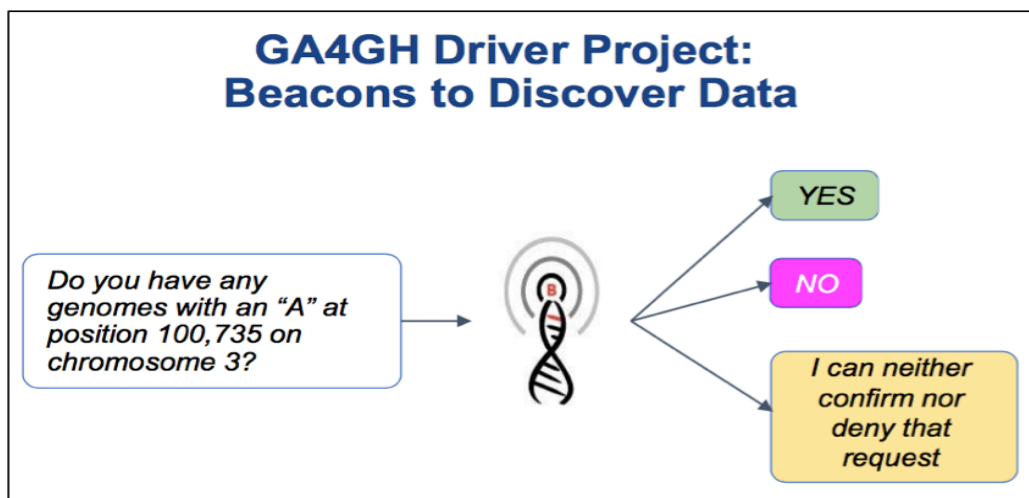
- ・ 子供のインフォームドコンセントについて。今後、小児がんや出生前診断についても検討。
- ・ 家族歴を集めることが重要。それにあたっての IC をどうすべきか。
- ・ Machine readable Consent の検討・開発

・ e-health のコンセント（誰がデータを見ることができ、誰が Edit できるか等の検討）

・ データの共有に関与する様々な Stakeholder 間の説明責任について（研究者、研究参加者、Journal、企業、倫理委員会、データベース・Biobank 等）>Task Team を立ち上げて Accountability Policy のドラフトを作成した。

5 : Beacon Project

検索対象 DB が条件（ゲノム上の特定の位置の塩基が、指定した塩基か否か）を満たすゲノムデータ（頻度だけでなく個人ゲノム）を持っているかを yes/no で返す。分散しているゲノムデータを検索しやすくすることが目的。



<https://ega.crg.eu/beacon/#/>

EGA の例

EGA Beacon EMBL-EBI CRG

Username Password Log in

Query available datasets

Select a dataset: All GRCh37 (87247762 variants)

Reference genome: GRCh37

Chromosome: 1

Position: 0

Allele: A

Find

Note: The EGA archives a large number of datasets, some of which are **Public** available. If you have an account on this website, you can access to **Controlled** datasets. To access **Controlled** datasets can be done by contacting the relevant Data Access Committee (DAC), whose details are displayed on the Dataset description page under "Who controls access to this dataset" (click on the Dataset ID to go to the Dataset page). Once you have access to a dataset, you will be able to query it at EGA Beacon.

EGA ID	Short title	Access type
EGAD00001000...	Low-coverage whole genome sequencing; variant calling, genotype calli...	PUBLIC
EGAD00001000...	This sample set comprises cases of schizophrenia with additional cogniti...	CONTROLLED
EGAD00001000...	This sample set of UK origin consists of clinically identified subjects with ...	CONTROLLED
EGAD00001000...	The sample selection consists of subjects with schizophrenia (SZ), autis...	CONTROLLED
EGAD00001000...	The MGAS (Molecular Genetics of Autism Study) samples are from a clini...	CONTROLLED
EGAD00001000...	Two groups of samples with diagnosis of schizophrenia or schizoaffectiv...	CONTROLLED
EGAD00001000...	The BioNED (Biomarkers for Childhood onset neuropsychiatric disorders)...	CONTROLLED
EGAD00001000...	The Tampere Autism sample set consists of samples from Finnish subjec...	CONTROLLED
EGAD00001000...	The entire sample collection consists of 2756 individuals from 458 familie...	CONTROLLED
EGAD00001000...	Samples from three sources: the Genetics and Psychosis (GAP) set cons...	CONTROLLED
EGAD00001000...	Low-coverage whole genome sequencing; variant calling, genotype calli...	PUBLIC
EGAD00000000...	Procardis study for coronary artery disease. GWAS study. 3352 cases - ...	REGISTERED
EGAD00010000...	Summary statistics from Haemgen RBC GWAS (Anemia)	REGISTERED
EGAD00000000...	Aggregate results from a case-control study on stroke and ischemic stro...	REGISTERED
EGAD00000000...	WNT-signaling and Dupuytren's Disease. GWAS analysis. 856 cases - 28...	CONTROLLED
EGAD00001000...	These samples are a subset of a nationwide collection of Finnish autism ...	CONTROLLED
EGAD00001000...	These Finnish schizophrenia samples have been collected from a popula...	CONTROLLED
EGAD00001000...	These affected schizophrenia families have been diagnosed using the SA...	CONTROLLED
EGAD00001000...	This is an Irish sample set of individuals with ASD (approximately 50% wi...	CONTROLLED
EGAD00001000...	This sample set consists of subjects with schizophrenia recruited from p...	CONTROLLED

Total Items: 22

Yes データがどのスタディ（データセット）に含まれるか、その集団における頻度を示す。

1 年間で 15 組織で beacon が立ち上がり、252 のデータセットが提供されている(2015/6/1 現在)。
<http://ga4gh.org/#/beacon> >>NBDC にも Beacon への参加を期待している。