

○笹沼俊一¹、大波純一¹、村田武英¹、三輪佳宏¹

¹. 国立研究開発法人理化学研究所 バイオリソース研究センター 遺伝子材料開発室

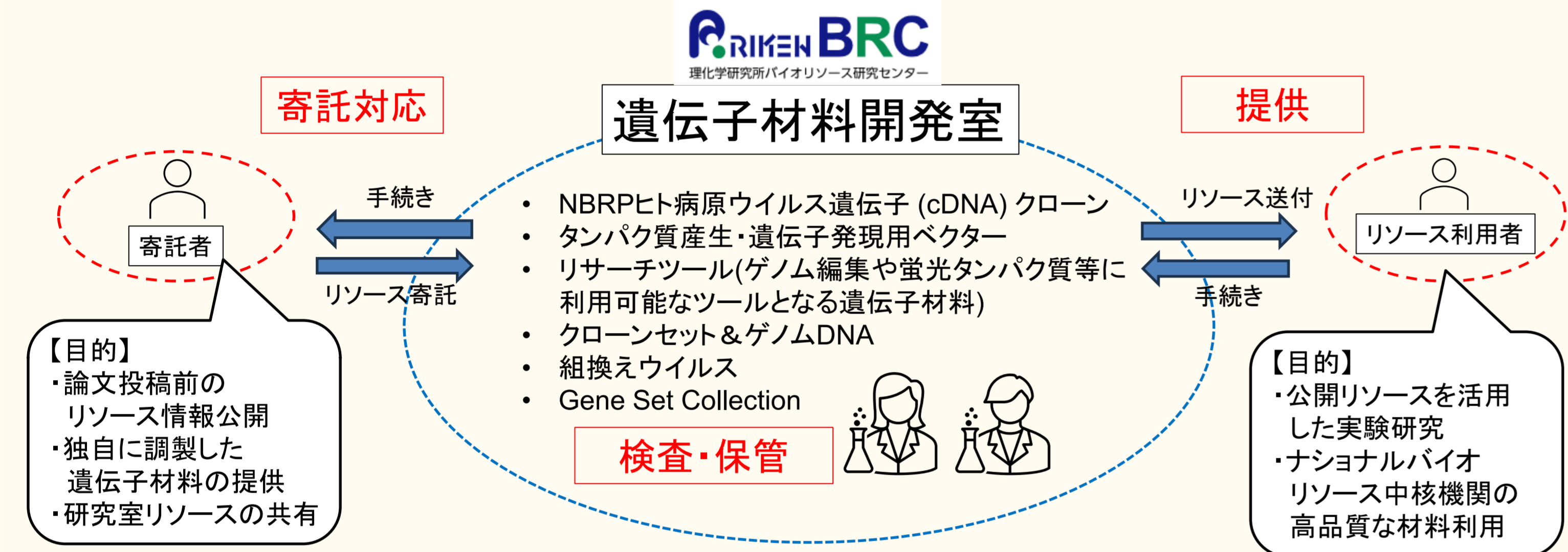
要旨

理化学研究所バイオリソース研究センター(BRC)の遺伝子材料開発室では、文部科学省ナショナルバイオリソースプロジェクト(NBRP)「遺伝子材料」の中核機関として、リソースの提供事業を行っている。遺伝子材料リソースには、各種生物のゲノムライブラリや可視化レポーターとして利用可能な発現ベクター、ゲノム編集及び遺伝子導入用プラスミドクローン等が含まれる。これらの多くは寄託者が何らかのプラスミドから独自に遺伝子をカスタマイズし作製されている。登録されたバイオリソースは、サンプルごとに利用目的や培養条件などのメタデータが付加されて構造化データとして収集されている。

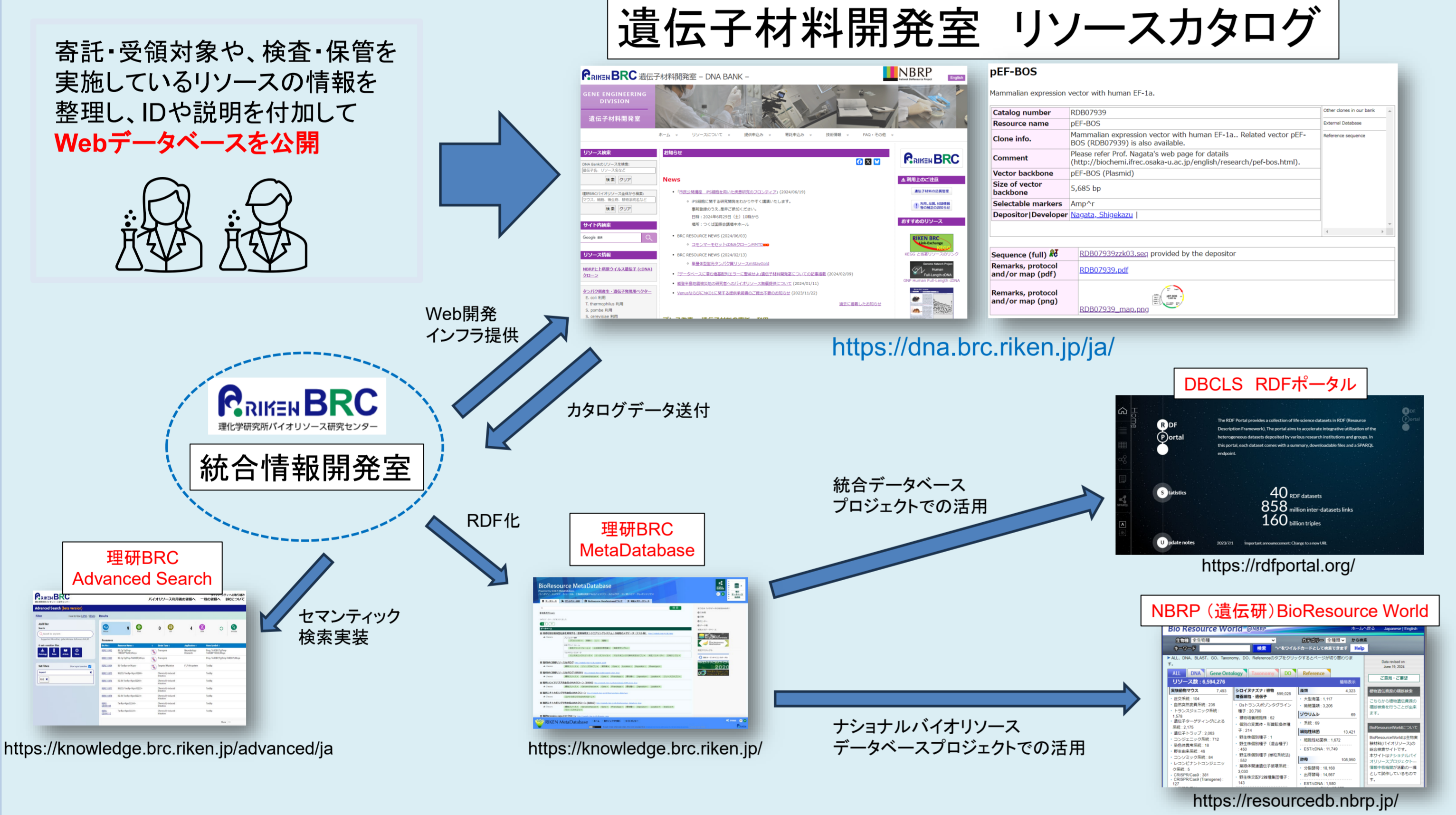
今回この構造化データの中で由来ベクターを示す「Backbone情報」を利用して、関係性グラフデータを構築した。これによって作製ベクターの由来となりやすいリソースを新たに分析できた他、数世代に渡りBackboneとして使われる、カスタマイズし易いベクター情報を抽出することができた。この情報を用いて、Web上のカタログデータの表示や検索アルゴリズムを刷新し、バイオリソース利用者にとってより利便性の高いサービスを提供していく。

理研BRC遺伝子材料開発室の事業

遺伝子材料リソースの「寄託対応」と「提供」、「検査・保管」



カタログデータの構築と公開



リソースとカタログデータベースの、さらなる高品質化を目的として現況の調査と分析を実施

ベクターのBackbone関係

遺伝子材料開発室 リソースカタログのメタデータ1件ごとに「Vector backbone」という枠が存在し、カスタマイズされた新規ベクターには、製作の元となったベクター情報が1件登録されている(下図点線枠)。

pCAG-EGxxFP	
Validation plasmid for the effectiveness of guide RNA in CRISPR/Cas9 system.	
Catalog number	RDB13532
Resource name	pCAG-EGxxFP
Clone info.	Validation plasmid for the effectiveness of guide RNA in CRISPR/Cas9 system.
Comment	For validation of sgRNA on double strand breaks by observation of green fluorescence. MCS: BamHI, NheI, PstI, SalI, EcoRI, and EcoRV.
Vector backbone	pCAGGS (Plasmid)
Selectable markers	Amp ^r
Gene/insert name	Aequorea victoria GFP cDNA
Depositor/Developer	Ikawa, Masahito

よって、ベクターごとにbackboneを辿ることでその系譜を構築することができると考えられた。また、多数のベクターのbackboneとなっている対象を判別することができれば、その度合いを比較することでベクターのインパクト評価もできることが予想された。これらを目的とし、表形式DBであったリソースカタログのグラフデータ化と関係性ネットワーク分析を実施した。

Backbone関係での系譜構築とインパクト評価のために、リソースカタログのグラフデータ化とネットワーク分析

分析手順

カタログデータの処理とグラフデータ化

カタログデータは2024年7月17日時点におけるマスターデータを利用した。処理のプロセスを以下に示す。



構築したグラフとネットワーク分析

ベクターの系譜と登録者の構成情報

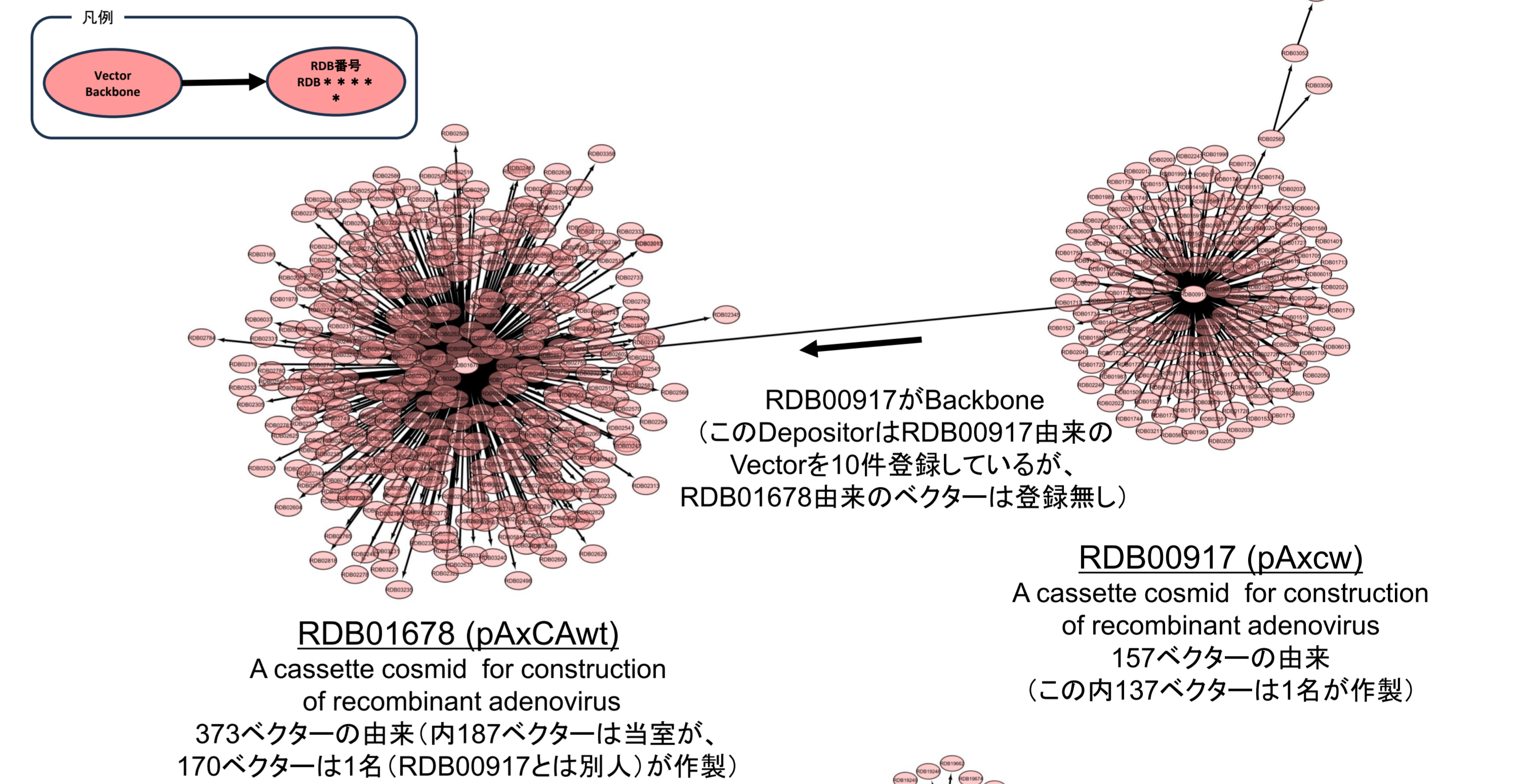
ベクター同士のBackbone関係の二項情報1685件から、最大524ベクターが繋がった117件の集合を得た(右図・右表)。これに關係するRDB番号ごとのDepositor(登録者)は66人おり、当室も1Depositorとして含まれていた。二項情報1685件の中で、当室は618件を登録していた。239件や215件を一人で登録するDepositorが存在する一方、1件だけ登録しているDepositorも15人存在していた。

このことから、カタログのベクター作製は必ずしも少人数だけで実現しているわけではなく、ある程度の多様な研究者によってなされていると判断した。

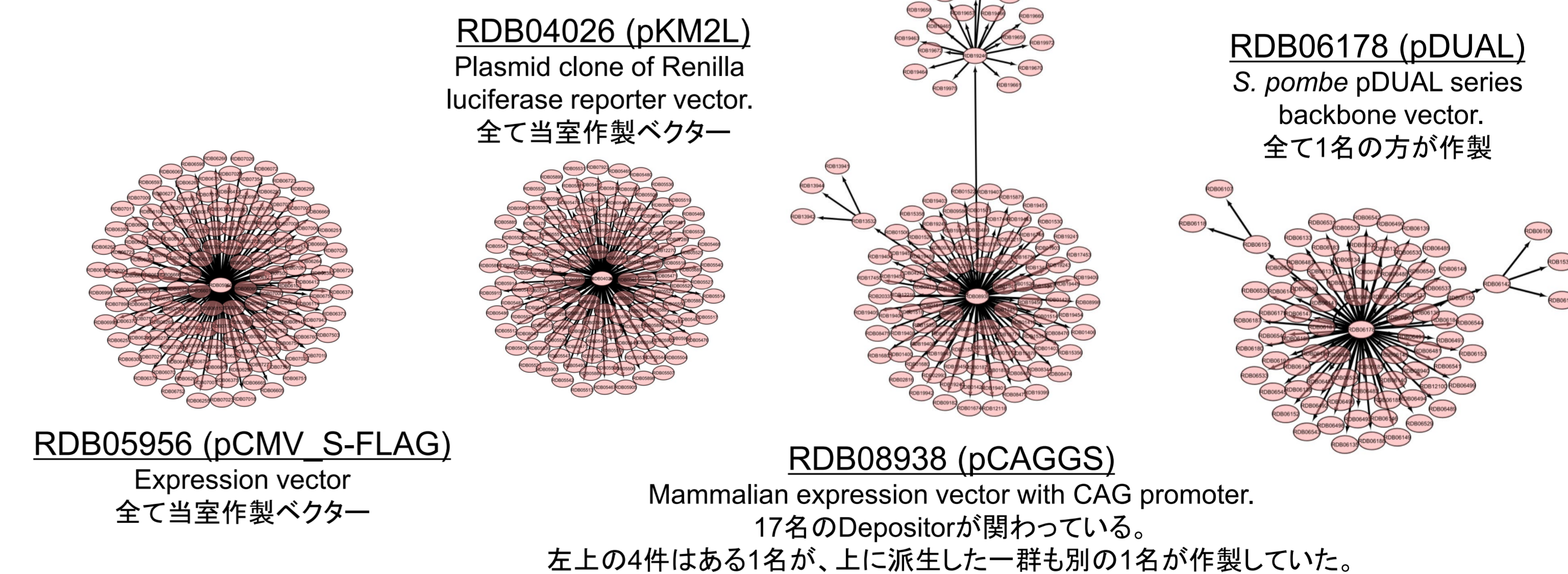
集合内のベクター数	集合数
1(自身を由来とする)	41
2	23
3	11
4	7
6	2
7	3
8~	30

Backboneベクターとなりやすい集合の分析

凡例に従い、Backbone関係のベクターを矢印で繋ぎ表示した。派生するベクター数が最も多かったのはRDB00917由来の集団で、派生したRDB01678から更に多くのベクターが作製されていた。



接続数が2番目から5番目に多い集団を以下に示す。



議論・今後の展望

バイオリソースのカタログデータとして登録されているベクター同士のBackbone関係を元に、多くのベクター作成の元となっているベクター集団の上位5件を分析した。その結果、(1) 1人のDepositorが全てのベクターを作製している場合と、(2) 多人数が1つのベクターから複数のベクターを作製している場合の2種類の集団が確認された。これらの違いは、新しいベクターを作製する場面や元ベクターを選定する判断理由が影響すると考えられる。最も多くの派生ベクターの由来となっていたのは、「RDB00917 (pAxcw) A cassette cosmid for construction of recombinant adenovirus」であった。この原著論文 (PMID: 7479022) も引用数729回とインパクトのある結果であり、組換えアデノウイルスツールとしてこのカタログ情報が広く認識された結果であると予想される。引き続きBackbone以外の情報も含めベクターの評価に繋がる指標を検討し、外部検索エンジン用の埋め込みメタデータや、Webカタログの推薦機能の評価値の一つとして採用を検討していく。

利益相反

本発表に関連し、開示すべき利益相反(COI)関係にある企業・法人組織や営利を目的した団体はありません。