

課題番号 : 28指1301

研究課題名 : バイオリソースからの情報を扱う臨床研究の品質管理方法およびシステムの開発研究

主任研究者名 : 田中紀子

分担研究者名 :

キーワード : omics; data management; database; clinical bioinformatics; standardization

研究成果 :

平成 28 年度の研究成果報告

初年度は前年度までにある程度完成されていたシステムを所内公開へ向けて見直すために、以下の 4 つの項目についてそれぞれ検討を行った。

#### 1. 計算環境設計の見直し

本研究計画時点では、ネットワーク環境が JCRAC およびセンター全体で大きく変わりつつあるので、再度サーバーをどこに設置し、どのような形式で接続するのが、現在の指針およびネットワークセキュリティ対策技術に最も対応できるのか、JCRAC データセンターだけではなく、病院の医療情報部、バイオバンク科などに協力をお願いし、再検討する必要性があった。また、国立循環器病センターではすでに検体からの情報もクラウドサーバーを利用して解析する体制を整えていると伺っているので、当センターでもクラウド利用で問題ないと判断された場合は、アマゾンなどのクラウドサービス利用も検討することとしていた。しかし、その後国府台以外で Office365 が導入されたことや、センター内のネットワーク環境が整ってきたことを受け、現在の JCRAC データセンターのネットワーク環境から病院ネットワークの DMZ 領域にサーバーを移動することを計画した。

#### 2. DNA メチル化測定用ツールの見直し

テスト環境が研究所から Galaxy サーバーへのアクセスが可能となった場合、実際に研究所からサーバーへアクセスしてもらい、現在進行中のプロジェクトなどで実験データの管理に使っていただき、改善が必要な項目について調査を行う計画であったが、所内ネットワーク環境のプロキシ設定が厳しくなったことから、ssh 接続および VPN からの接続は不可能となったため、サーバーの設定を WWW 上でだけでも NCGM ギャラクシーへアクセス可能な状態に一旦する必要があった。そこで、テストであっても実データはサーバーへ上げないこととして、出版済みデータとダミーデータのみでサーバーへの接続テストやアルゴリズムの確認を行うのみにとどまり、新規データなどによる十分な見直しは今年度へ持ち越すこととなった。

#### 3. SNP アレイ用ツールの開発

研究協力者から現在使用中の SNP アレイについて調査を行い、どのようなツールが必要であるか調査の上決定は計画通り行われた。現在は決定されたツールの組み込みデザインを設計し、プログラミングに着手している。

#### 4. 電子実験ノートの機能の拡張

当初 JAVA で開発された実験ノートの DNA メチル化データについての入力項目の見直しを行い、実際に研究所の先生に使っていただき改善点を洗い出し、改善を行った。それに伴い、使い勝手が悪かったので、開発言語を Ruby に変更し新たに作り直し、提供を行い、テスト使用をお願いしている。また、口頭試問より大筋のアプリケーションデザインに変更がないことをみこして、SNP アレイ測定のための機能を追加できるような手順を整えた。

Subject No. : 28 指 130 1

Title : System development based on the study for quality control in clinical bioinformatics research

Researchers : Noriko Tanaka

Key word : omics; data management; database; clinical bioinformatics; standardization

Abstract :

During last three years, I and my colleagues developed a new system to support for managing quality in clinical bioinformatics research, which can deal only with DNA methylation data.

During three years from last year, we planned to continue developing our system to add functions for Genome Wide Association Studies (GWAS) and open it to NCGM faculties.

In the last year, we considered the following each part of the system:

1. Reconsider computing environment

At the project planning phase, the network environment in our institute had been changing and we needed to plan to discuss which place (or network) was the best for the server to be placed for open it to our communities. Finally, we decided to place the server in the physical subnetwork (DMZ) of our organization's network.

2. Revise the tool for DNA methylation data

Because the researchers inside the hospital could not access the server system due to the proxy settings managed by the hospital, we needed to test the system only with the opened or published DNA methylation data, and could not complete the test with unpublished newly corrected data. We will try to test the system with unpublished data after the server is moved to DMZ on NCGM network.

3. Develop new tools for SNP array data

According to the project plan, we examined which arrays were frequently used in Japanese genetic research communities and which quality control tools were needed for such common microarray platforms.

4. Extend the function in our online application for experimental management

During last three years, a web-based experiment notes to control the quality of experiments for epigenomic studies using Ruby on Rails scripts. In the last year, we tested the developed application and finalized it for open. We created the documents of operation procedures for the person in charge of development, and made arrangement to add another function for measurements from SNP microarray platforms.

研究目的:

バイオリソースからの情報を扱う臨床研究の品質管理方法の検討およびシステムの構築

Specific Aim 1:

セキュアな環境下で多層かつ異質的分子生物学的情報を格納するストレージシステムとGalaxyサーバーシステムとを連携したネットワークシステムの構築

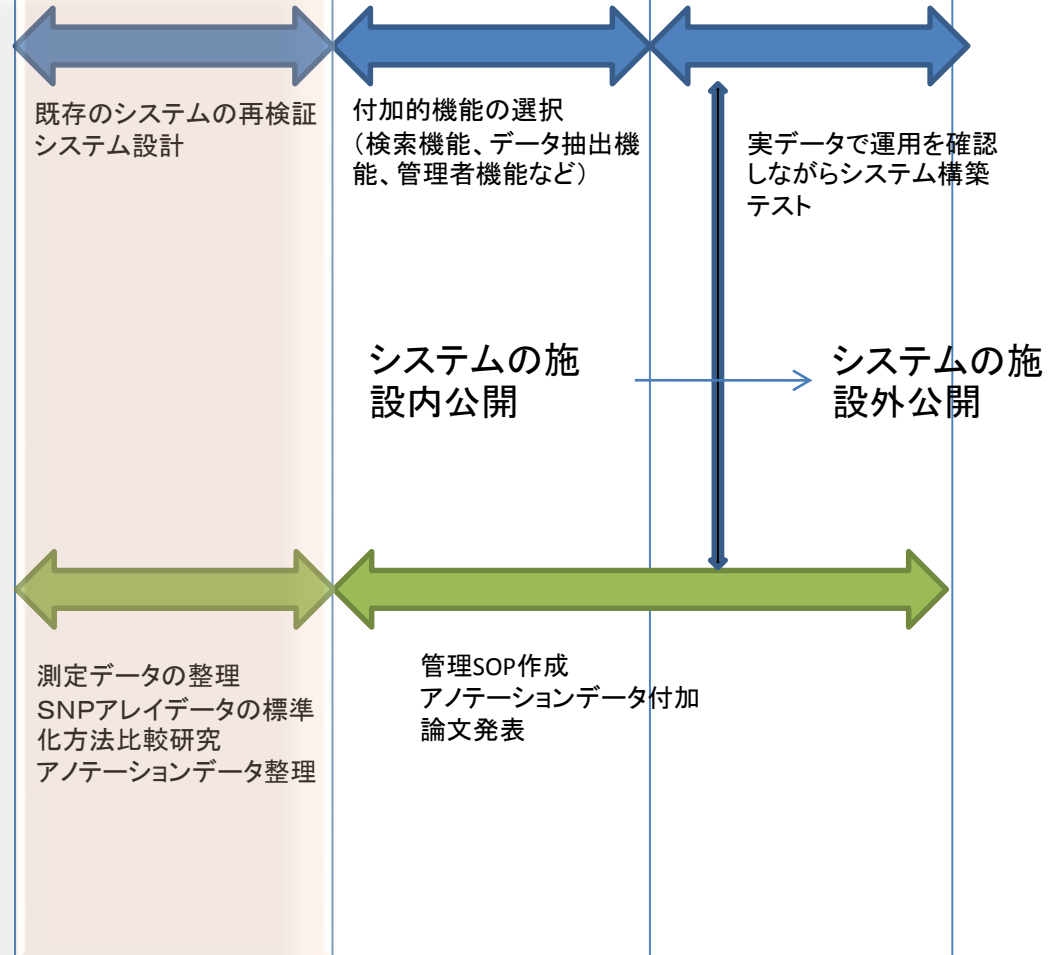
Specific Aim 2:

SNPアレイデータ(実データ)の品質管理工程の確立

平成28年度

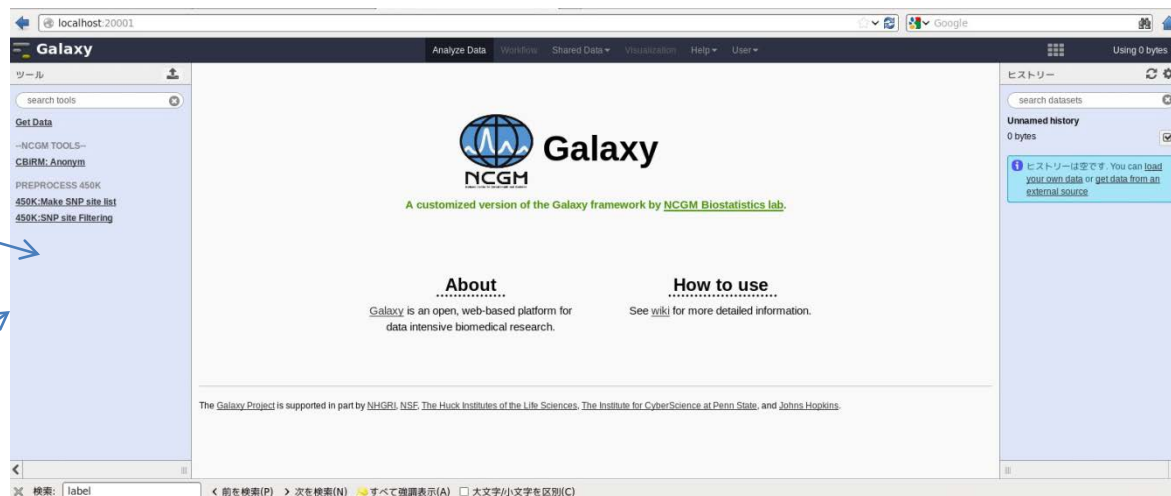
平成29年度

平成30年度



# 研究データの発生から公開までの一連の流れをテスト

SNP解析ツールの追加開発予定



実験者による実験に関するデータ入力

Galaxyでデータ加工・データ解析を実行

結核菌ゲノムの変異解析 編集履歴  
(作成者: 菅岡理人, 編集日時: 2016-09-19 13:39)

PROTOCOLS

growth protocol	00000
treatment protocol	00000
extract protocol	00000
label protocol	00000
hyb protocol	00000
scan protocol	00000
data processing	00000
matrix processed value definition	00000000
# matrix signal intensities value definition	000000

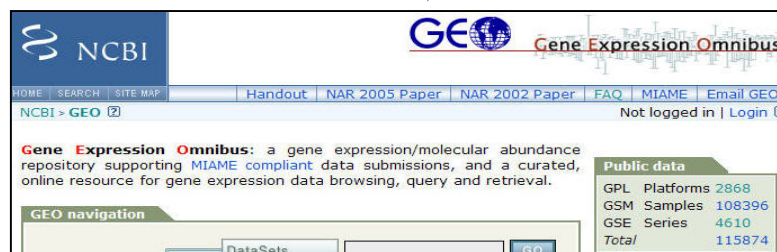
Free Form

SAMPLES

Sample name	title	source name	organism	sex	molecule	label	description	platform	free form 1	free form 2
0000	0000	0000	00000	00000	0000	0000	0000	0000	00000000	0000
0000	0000	0000	00000	00000	0000	0000	0000	0000	00000000	0000
0000	0000	0000	00000	00000	0000	0000	0000	0000	00000000	0000
0000	0000	0000	00000	00000	0000	0000	0000	0000	00000000	0000
0000	0000	0000	00000	00000	0000	0000	0000	0000	00000000	0000
0000	0000	0000	00000	00000	0000	0000	0000	0000	00000000	0000
0000	0000	0000	00000	00000	0000	0000	0000	0000	00000000	0000
0000	0000	0000	00000	00000	0000	0000	0000	0000	00000000	0000

公開用データ

A	B	C	D	E	F	G	H	I	J	K
SERIES										
title	test									
summary	test									
overall design	testoveralldesign									
contributor1	contributor1									
SAMPLES										
Sample name	title	source name	organism	characteristics: gender	characteristics: sample type	characteristics: disease state	molecule	label	description	platform
0	genomic DNA from cervical smear 1	cervical smear 1	Homo sapiens	Female	cervical smear	normal	genomic DNA	Cy6 and Cy3	normal cervical sample	GPL8490
111	genomic DNA from cervical smear 2	cervical smear 2	Homo sapiens	Female	cervical smear	normal	genomic DNA	Cy5 and Cy3	normal cervical sample	GPL8490
PROTOCOLS										
Sample name	title	source name	organism	characteristics: gender	characteristics: sample type	characteristics: disease state	molecule	label	description	platform
0	genomic DNA from cervical smear 1	cervical smear 1	Homo sapiens	Female	cervical smear	normal	genomic DNA	Cy5 and Cy3	normal cervical sample	GPL8490
111	genomic DNA from cervical smear 2	cervical smear 2	Homo sapiens	Female	cervical smear	normal	genomic DNA	Cy5 and Cy3	normal cervical sample	GPL8490



## 研究発表及び特許取得報告について

課題番号： 28指1301

研究課題名： バイオリソースからの情報を扱う臨床研究の品質管理方法およびシステムの開発研究

主任研究者名： 田中紀子

### 論文発表

論文タイトル	著者	掲載誌	掲載号	年
該当なし				

### 学会発表

タイトル	発表者	学会名	場所	年月
該当なし				

### その他発表(雑誌、テレビ、ラジオ等)

タイトル	発表者	発表先	場所	年月日
該当なし				

特許取得状況について ※出願申請中のものは( )記載のこと。

発明名称	登録番号	特許権者(申請者) (共願は全記載)	登録日(申請日)	出願国
該当なし				

※該当がない項目の欄には「該当なし」と記載のこと。

※主任研究者が班全員分の内容を記載のこ