

## ゲノムネットワークプラットフォームの構築 (詳細)

## ① ゲノムネットワークパイプラインの改良と運用

## [現状]

既存の生体分子間ネットワーク等に関する情報を標準化し、ゲノムネットワークプロジェクト参加機関において生成される様々な情報とともに蓄積された情報を統合し、公開するためのゲノムネットワーク統合データベースを開発・構築した。

## [今年度計画]

タンパク質などの生体分子の相互作用に係る情報の大規模データを統合・データベース化するための高速システム、「情報処理パイプライン」の運用をおこなうため、既存の生体分子間ネットワーク等に関する情報を標準化し、ゲノムネットワークプロジェクト参加機関において生成される様々な情報とともに集約・公開するためのプロトコル等を開発する。また蓄積された情報を統合し、公開するためのゲノムネットワーク統合データベースを開発・構築する。今年度は、横軸よりの PPI, DPI データの取り込みを中心にデータベースバージョン 2 の構築公開を進める。

## ②ゲノムネットワーク利用システムの開発

## [現状]

ゲノムネットワークプロジェクト内外に公開するためのゲノムネットワーク公開インターフェースの開発を行いコンソーシアム、一般向けの両方に関して公開した。

## [今年度計画]

情報処理パイプラインの生成するデータベースよりネットワーク情報を抽出・編集し、ゲノムネットワークプロジェクト内外に公開するためのゲノムネットワーク公開インターフェースの開発・公開を行う。また、本研究課題実施のために必要な技術等の開発も行う。今年度は、テキストマイニング機能の追加や DPI データなど新しいタイプのデータに対応した利用システムバージョンの公開を進める。

## ③ゲノムネットワークアルゴリズム等の研究

## [現状]

ゲノムネットワークプラットフォームに要求される要素技術として、ゲノムネットワークデータ検索アルゴリズムの研究、各種生体分子情報からのネットワーク情報の抽出アルゴリズムの研究とその表示に関する研究をおこない、実装に向けての試験を行った。

## [今年度計画]

ゲノムネットワークデータ検索アルゴリズムの研究、各種生体分子情報からのネットワーク情報の抽出アルゴリズムの研究とその表示に関する研究をおこなうとともに、各種統計解析結果をデータベースに統合し公開できるようにしていく。

## ④生命科学文献からの情報抽出とテキストマイニング技術の開発

## [現状]

公開データベース、論文等の文献情報からゲノムネットワークの解明に有用な情報を抽出・活用するため、専門用語の言語処理、テキストマイニングに関する技術開発を進め、PPIに関する知識情報抽出システムの開発を行った。

## [今年度計画]

公開データベース、論文等の文献情報からゲノムネットワークの解明に有用な情報を抽出・活用するため、専門用語の言語処理、テキストマイニングに関する技術開発等を行うとともに、プラットフォームとの連携を図り、解析データを反映するようにする。

#### ⑤蛋白質相互作用の変化を推定するアルゴリズムの開発

[現状]

選択的スプライシングによる蛋白質相互作用の変化を解析するため、選択的スプライシングに関するデータの収集・解析を進め、データベースの作成をおこなった。

[今年度計画]

ゲノムネットワークを解明するために、遺伝子発現産物であるタンパク質の多様性がどのような役割を果たしているのかという観点から、選択的スプライシングによるバリエーションの産生が招く蛋白質相互作用の変化と遺伝子発現ネットワークや代謝系ネットワークの制御のメカニズムとの関連を解析する。この過程で生み出されるアルゴリズムを1連のパッケージとして広く公開するとともに、プラットフォームとの統合をはかる。

#### ⑥垂直統合型細胞内分子相互作用ネットワーク分析情報基盤技術の開発

[現状]

垂直統合型細胞内分子相互作用ネットワーク分析情報基盤技術の開発を進めた。

[今年度計画]

ゲノムネットワークの研究に必要となる多様な情報を記述・表現し、様々なスケールで動的な分析を可能にする情報基盤を構築するために、システムバイオロジーマークアップ言語(SBML)、Cell Designer 分子相互作用エディターなどの技術開発とシステムの開発を行い、これらの多様な情報の記述と解析に対応させると共に、データの垂直統合を実際に行い、その有効性を実証する。