

# バイオ研究支援環境構築ソフトウェア BioINTEGRA

バイオ研究に必要なイントラネット環境を実現するバイオ研究支援環境構築ソフトウェアBioINTEGRAを開発した。

## 概要

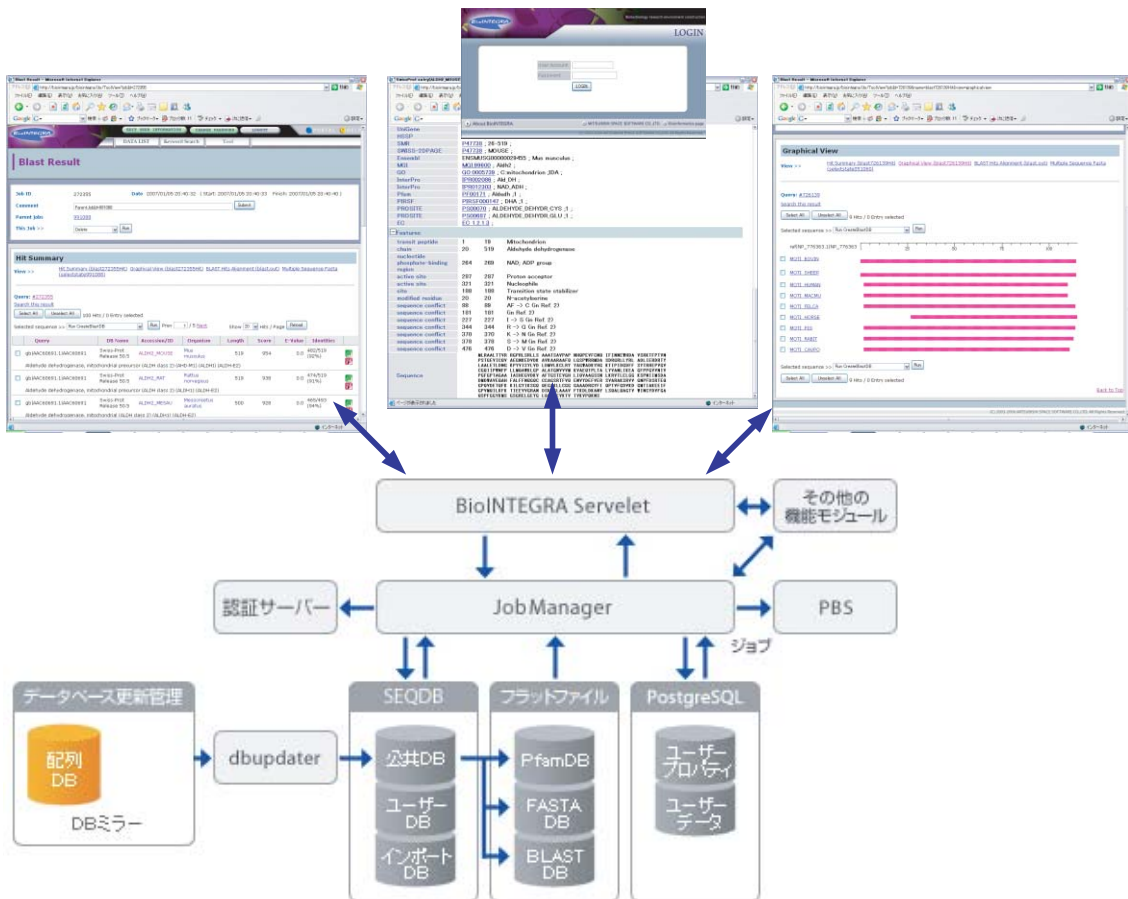
BioINTEGRAは、『統合』と『カスタマイズ』をコンセプトにした製品である。

BioINTEGRAは、PCクラスタシステムなどのHPC（ハイパフォーマンスコンピューティング）技術による高速処理環境を提供し、バイオ研究に必要な配列データベースの統合や配列解析ツールのWebサービスをイントラネット上に構築する。BioINTEGRAの導入により

バイオ研究者は導入作業に悩まされることなくバイオインフォマテックスの効能を満喫できるようになる。

また、BioINTEGRAはカスタマイズも得意である。パッケージ化されたオプション機能とカスタマイズサービスを組み合わせて多様なシステムを安価に構築できる。様々なニーズに対応したグループ共用システムとして最適なソリューションである。

BioINTEGRA Webサービス



## 開発の経緯

加速的に増大するデータベースや進化の著しい解析機器からのデータ分析に、時間が掛かる。そのためインフォマティクスを導入したいが、研究者にとっては必要なツール、データベース、計算機システムを選定、導入、

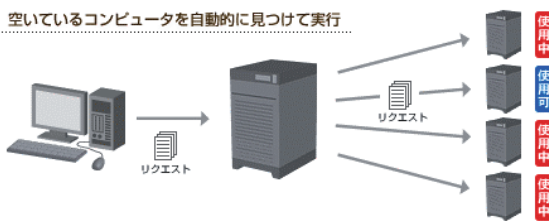
運用するには大きな負荷がかかり負担になっていることから、バイオ研究者に代わって必要なツール、データベース、計算機システムをインテグレートするソフトウェアを開発した。

## 特徴

- ・ 使用の規模に応じたスケラビリティ  
 数人規模で使用になるラボレベルの研究システムの場合は1台のPCサーバで稼働させることが可能。また、大規模なクラスシステムにインストールすることにより、ユーザ数百人を越えるセンターシステムを構築することができる。
- ・ 独自データベースエンジンSEQDBを採用した低価格設計  
 高度なインデックス作成機能を持ち、配列データベースの構築に適した独自データベースエンジンである"SEQDB"を採用することにより、高速検索と低価格設定を同時に実現した。
- ・ 様々な計算機リソースを共有  
 ユーザごとにアカウントとプロパティ（メールアドレスなど）を設定し、計算機リソース（CPU、メモリ、HDD）を共有することができる。
- ・ Webブラウザからのツール実行とパイプライン処理  
 Webブラウザからインタラクティブに操作できる“同期実行”のほかに、バッチ処理のような“非同期実

- 行”をサポート。（ジョブの実行状況はイベントごとにメール通知）
- ・ 代表的なバイオインフォマティクスツールをプリインストール  
 BLAST、ClustalW、HMMERをプリインストール。さらに、別売の自動ダウンロードツールDr.Mirrorと接続することで、解析に必要な公共データベースの更新～インデックス作成までの自動化が可能。最新情報の維持に無駄なコストがかからない。
- ・ インデックスで補完して、キーワードで絞りこみ  
 “キーワードのフルスペルが思い出せない” そんなときには、インデックス機能がとても便利。検索フィールド（OrganismやDefinition）ごとに、キーワードを補完。キーワード検索の結果を検索フィールドで絞り込みたい場合に大変便利。
- ・ 複数のツールを続けて実行できる“パイプライン処理”を実現  
 すべてのツールはWebブラウザでの操作性を統一。さらに、BLASTの実行結果をClustalWの入力にするなど、解析ツールの結果を次の解析ツールの入力として“パイプライン的に”実行することができ、わずらわしいフォーマット変換やカット&ペーストが不要。
- ・ オリジナルのBLASTデータベースを簡単作成  
 キーワード検索の結果やBLAST検索の結果など、検索や解析の結果得られた配列群を処理しオリジナルのBLASTデータベースを作成する”カスタムDB作成機能”を実現。顧客の研究ニーズに合った独自BLASTデータベースを簡単に作成できる。

### 空いているコンピュータを自動的に見つけて実行



### インデックス検索～キーワードによる絞り込み



### オリジナルのBLASTデータベース～キーワードによる絞り込み



## siRNA配列設計システム SISNIPER

従来の設計手法だけでなく、最新の研究成果（OffTargetEffectなど）を取り込み、さらに実験結果をフィードバックすることで高い活性を持ったsiRNA配列を高精度で設計することができるsiRNA配列設計システムSISNIPERを開発した。

### 概要

・本格設計

今までのsiRNA配列設計手法は、生化学者の経験則から生まれたものがほとんどで、的中精度が低いという問題点があった。SISNIPERは確率・統計学、バイオインフォマティクス、分子生物学のそれぞれの専門スタッフが知識を結集して作成したアルゴリズムを利用しており、これが、高い的中率を生み出している理由である。

・実験データのフィードバック

ジェノメディア（株）が保有するハイスループット解

析技術を駆使して得られたデータはシステムに取り入れられ、設計ロジックに反映される仕組みになっているため、SISNIPERは常に進化し続ける。

・設計を中心としたサービス

我々のコアコンピタンスはsiRNA設計技術であり、お客様に活性のあるsiRNAをご提供することが至上命題であるが、SISNIPERを核としたシステムを構築することで、siRNAを使った研究を総合的に支援するようなパッケージ提案も可能である。

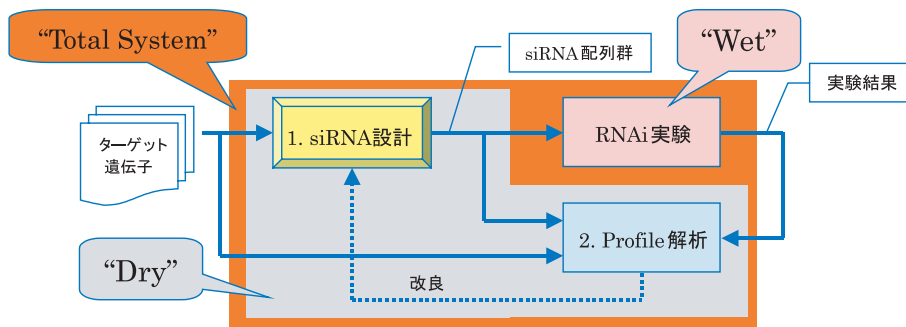


図1 システムの概要

### 開発の経緯

大阪大学、産総研、阪大初ベンチャー企業であるアンジェスMGが出資した子会社のジェノメディア（株）と

共同で、地域新生コンソーシアム研究開発事業（平成15～16年度）の助成を受けて開発した。

### 特長

・高い予測精度

今までのsiRNA配列設計手法は、生化学者の経験則から生まれたものがほとんどだったため、的中精度が低いという問題点があった。SISNIPERは確率・統計学、バイオインフォマティクス、分子生物学のそれぞれの専門スタッフが知識を結集して作成したアルゴリズムを実装することで、高い的中率を実現した（図2参照）。

・実験データのフィードバック

設計精度を向上させるために、ジェノメディア（株）社が保有するハイスループット解析技術によって得られた実験データを利用している。実験データをフィードバックすればするほど、設計精度が向上される仕組みである。

・設計を中心としたサービス

SISNIPERとBioINTEGRAを連携することで、siRNAを使った研究を総合的に支援するようなシステムを構築することができる。

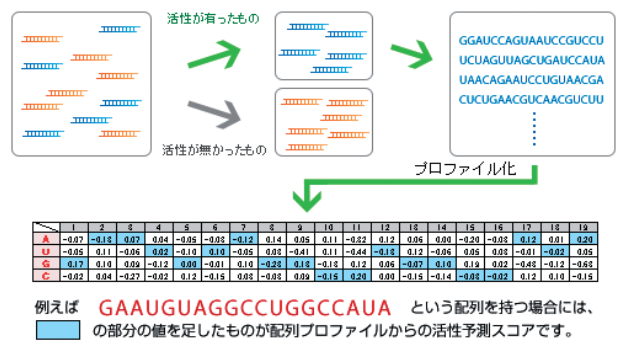


図2 予測精度を高めるためのアルゴリズム

## バイオ文献検索支援システム

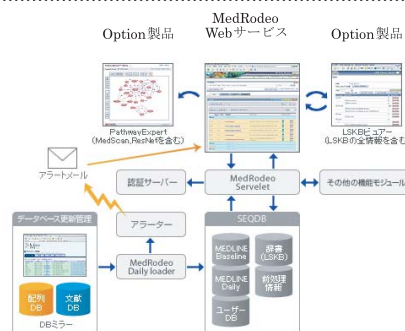
## MedRodeo

独自の高速文献・配列検索エンジンSEQDBを採用したバイオ文献検索支援システムMedRodeoを開発した。

## 概要

急速に進歩を続ける生命科学分野においては、タイムリーな情報獲得の重要性が増大しており、最新の研究情報と研究現場の情報ギャップを如何に解消するかが勝ち抜くために重要な課題となっている。MedRodeoは、最新の遺伝子辞書と高速文献検索エンジンを融合した極めて能力の高い知識獲得システムであり、「誰でも、簡単に」MEDLINEから「遺伝子」「疾患」「化合物」「組織」

といった重要なキーワードやそれらに関する文献を整理分別できる。



## 開発の経緯

BioINTEGRAを導入した顧客からのニーズにより開発を行い、BioINTEGRAのオプション機能として売り

出したが、MedRodeo単体のニーズが高いため、単品製品として開発を行い05年1月から販売開始した。

## 特長

1,500万件にも上る膨大な文献データベースであるMEDLINEから有用な知識を取り出すには、卓越した「検索能力」が必要とされる。しかし、MEDLINE用に公開されたインターフェースであるPubMedでは、残念ながら研究者個人の「検索能力」に強く依存してしまい、特に得意分野から少し外れた情報の検索では重要な情報を見逃す場合が少なくない。

MedRodeoは「誰でも、簡単に」MEDLINEから重要なキーワードを自動的に抽出するインターフェースを備えており、情報検索の初心者でも大量のMEDLINE文書から「遺伝子」「疾患」「化合物」「組織」という研究現場でも必要とするキーワードを瞬時に識別することができる。候補遺伝子の抽出や文献の分類など日々の活動で誰もが見落とし無くMEDLINE情報の整理・活用ができるようになる画期的なシステムである。

・増え続ける文献の絞込検索に圧倒的な威力を発揮

例えば、癌関連遺伝子として最近注目されているp27遺伝子をPubMedで検索した場合、ヒットする文献数は3,364件にも達し、埋もれた重要情報の探索には大変な

労力を要する。MedRodeoなら、簡単な操作で絞り込み条件を構成することができ、必要性に応じて2,078件から6件までの絞込み結果が得られる。

・使いやすいWebインターフェースを装備

直感的な操作で絞り込み検索条件を構成できるWebインターフェースを装備している。例えば、研究テーマをキーワードにして主要な因子の抽出を行いたい場合は、GDCT (Gene/Disease/Chemical/Tissue) 分類ビューが役に立つ。

・デイリーアップデート、E-Mailアラート機能

デイリーアップデートの後に、E-Mailアラート機能によって、キーワードに合致した最新文献をE-mailでお知らせする。実験テーマ/解析手法をキーワードにして実験に必要な参考文献を迅速に入手できる。

・さまざまなカスタマイズが可能

MedRodeoには、カスタムGUI、MGI等のデータベース、LSKB等の辞書の追加が可能。さらにユーザの要望に基づく仕様追加も可能。お仕着せではできないきめ細かな対応が可能。

## 関連製品

MedRodeoを応用した上位製品として創薬研究情報基盤システムBioElephantがある。

BioElephantは、配列情報、文献情報、実験データなどに関して創薬研究に必要とされる情報基盤を構築し、最新の研究情報を研究現場へスムーズに浸透させるソリューションである。BioElephantは、三菱スペース・ソフトウェア、ワールドフュージョン、米国ARIADNE

GENOMICSのバイオインフォマティクステクノロジーを結集した“創薬研究プラットフォーム”である。データメンテナンス等実績のあるテクノロジーを採用し、安定した動作を実現している。社内データと公共データの連携をスムーズに行うと同時に、社内実験データ、配列データと公共データとの連携も行うことができる。



## 遺伝子発現相互作用推定プログラム

# Minos

遺伝子発現データから、遺伝子相互作用ネットワークを推定するソフトウェアを開発した。

### 概要

Minos (Mathematical gene interaction Network Optimization Software) は、遺伝子・タンパク質発現量の時間変化データ (タイムコース・データ) に基づき、遺伝子・タンパク質間の相互作用関係 (促進、抑制) を

推定し、推定した遺伝子・タンパク質間の制御関係をパズウェイ図のように2次元及び3次元描画 (図1) により図示するソフトウェアである。

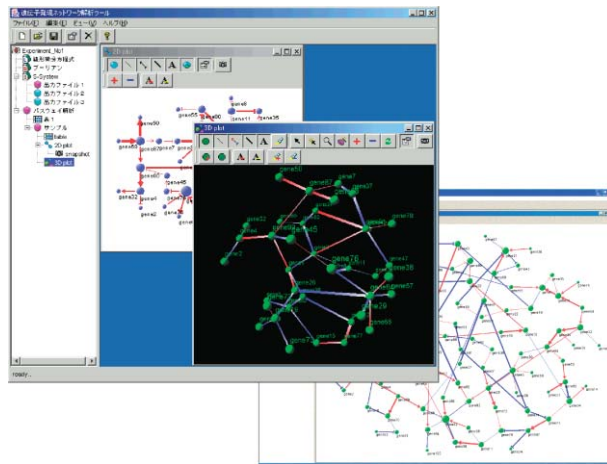


図1 二次元、三次元パズウェイ描画出力例

### 特長

Minosのコンセプトを図2に、以下に特徴を示す。

- ①マイクロアレイ・DNAチップ等の遺伝子発現解析のデータより遺伝子間相互作用を推定。
- ②遺伝子間相互作用を推定するソルバーとして3つのモデルを搭載。(図3) (S-Systemモデル、線形微分方程式

式モデル、ブーリアンネットワークモデル)

- ③GUIユーザインタフェース設計。(解析ソルバ、入力データ、出力結果等をプロジェクトとして管理)
- ④パズウェイ自動描画機能を用いて推定結果を視覚的に表示。

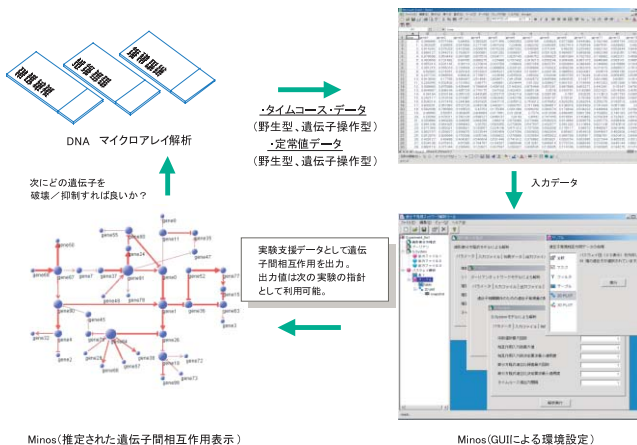
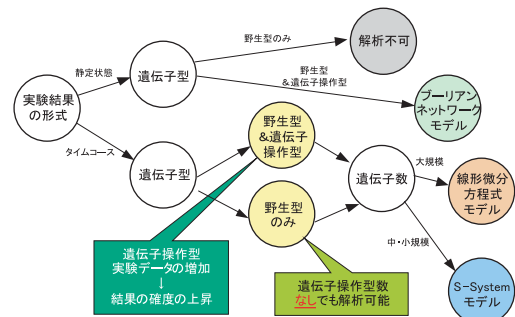


図2 Minosのコンセプト



\* Minosは、静的状態のデータを使用するブーリアン・ネットワーク・モデル、タイムコースデータを使用する線形微分方程式モデルとS-Systemモデルを、ソルバーとして持つ。

図3 ソルバーと実験結果形式の関係