

ライフサイエンス統合DBプロジェクト: ワークフローによるゲノム・分子情報解析

Life Science Database Integration Project :
Workflow Analysis for Genome and Molecular Information

●研究の目的

●ライフサイエンス研究分野では、その細分化や専門化に伴い、様々なデータベース、解析ツール、予測ソフトウェアが開発され膨大な数に達しており、これらを有効かつ効率よく利用するには様々な課題を克服する必要があります。生命情報工学研究センター(CBRC)では、文部科学省のライフサイエンス統合データベースプロジェクトに参画し、これら分散するソフトウェア解析やデータベース検索等を、一つ一つユーザが実行するのではなく、一連の処理の流れを定義することにより、効率的に短時間に実行するためのワークフローに関する技術開発を進め、統合化を目指した実用的な情報基盤開発に取り組んでいます。(図1)

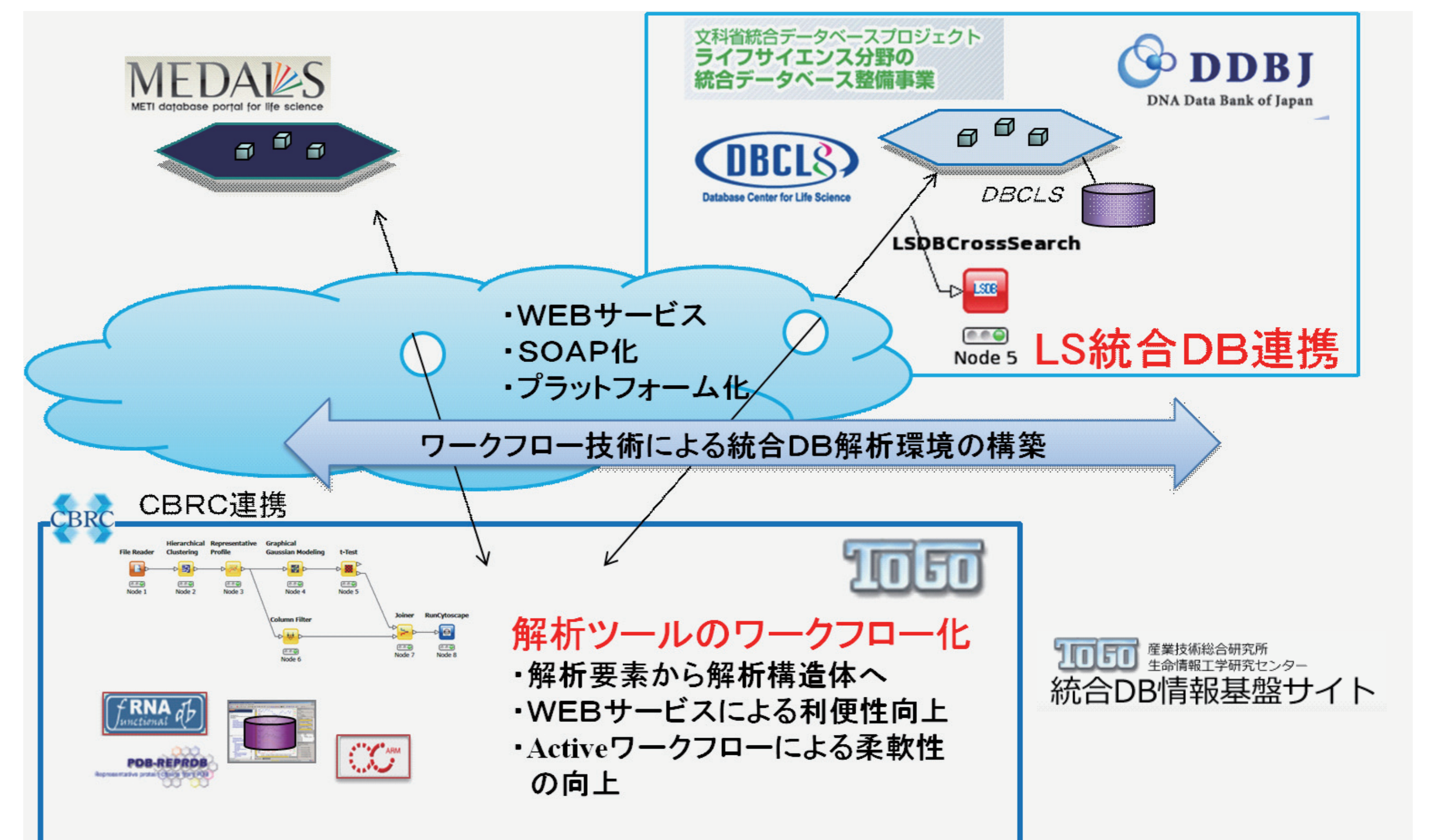


図1 情報統合プロジェクトにおける開発概要

●研究の成果

ここで開発していますワークフローとしては、WEBワークフローとアクティブ・ワークフローの二種類があります。

●WEBワークフロー

このワークフローは予め定義された処理を組み合わせ ており、ユーザは許された範囲内でのみパラメータ等の指定が可能で、実行が容易なWEBサービスです。(図2)

●アクティブ・ワークフロー

柔軟性を加味し、直観的なGUIを用いた操作性に優れたKNIMEとよばれるプラットフォーム環境を使用しています。このプラットフォーム上で、ユーザはノード化された解析要素であるツール、データベース、Webサービスとを自由に組み合わせることが可能であり、使いやすさと解析処理の流れを決定できる柔軟性を考慮したワークフローです。(図3)

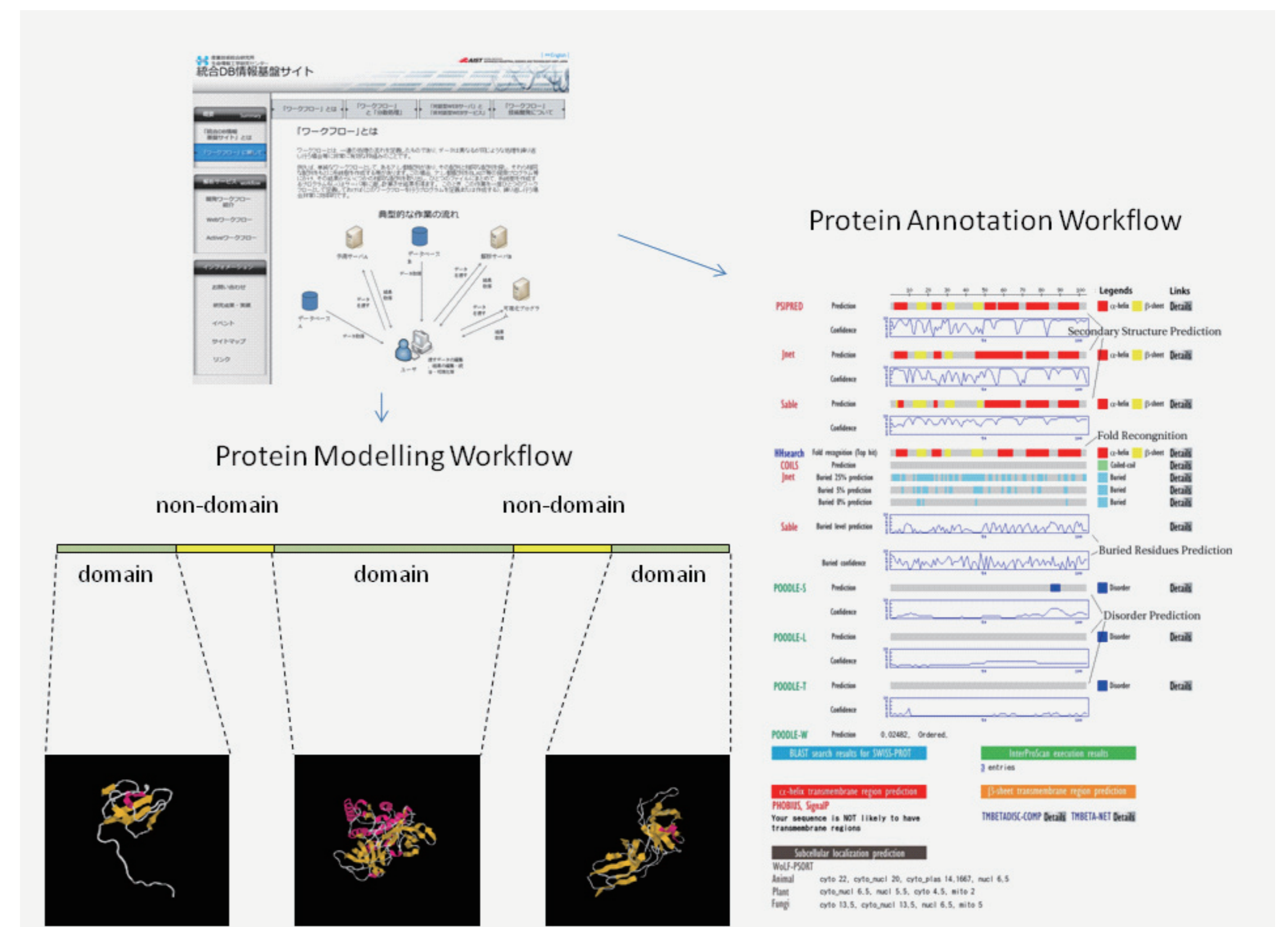


図2 WEBワークフロー

●将来の展望

●ライフサイエンス研究分野における専門化・細分化されたデータベースや解析ツールの迅速な変化や新規情報技術へ柔軟に対応できるワークフロー解析として期待できます。

●プラットフォームの利用によりデータベースと解析ツールを組み合わせたシームレスな大規模大量データ解析を可能とする技術開発へと繋がります。

●参考

アクティブ・ワークフローは統合DB情報基盤サイトからダウンロード可能です。
<http://togo.cbrc.jp>

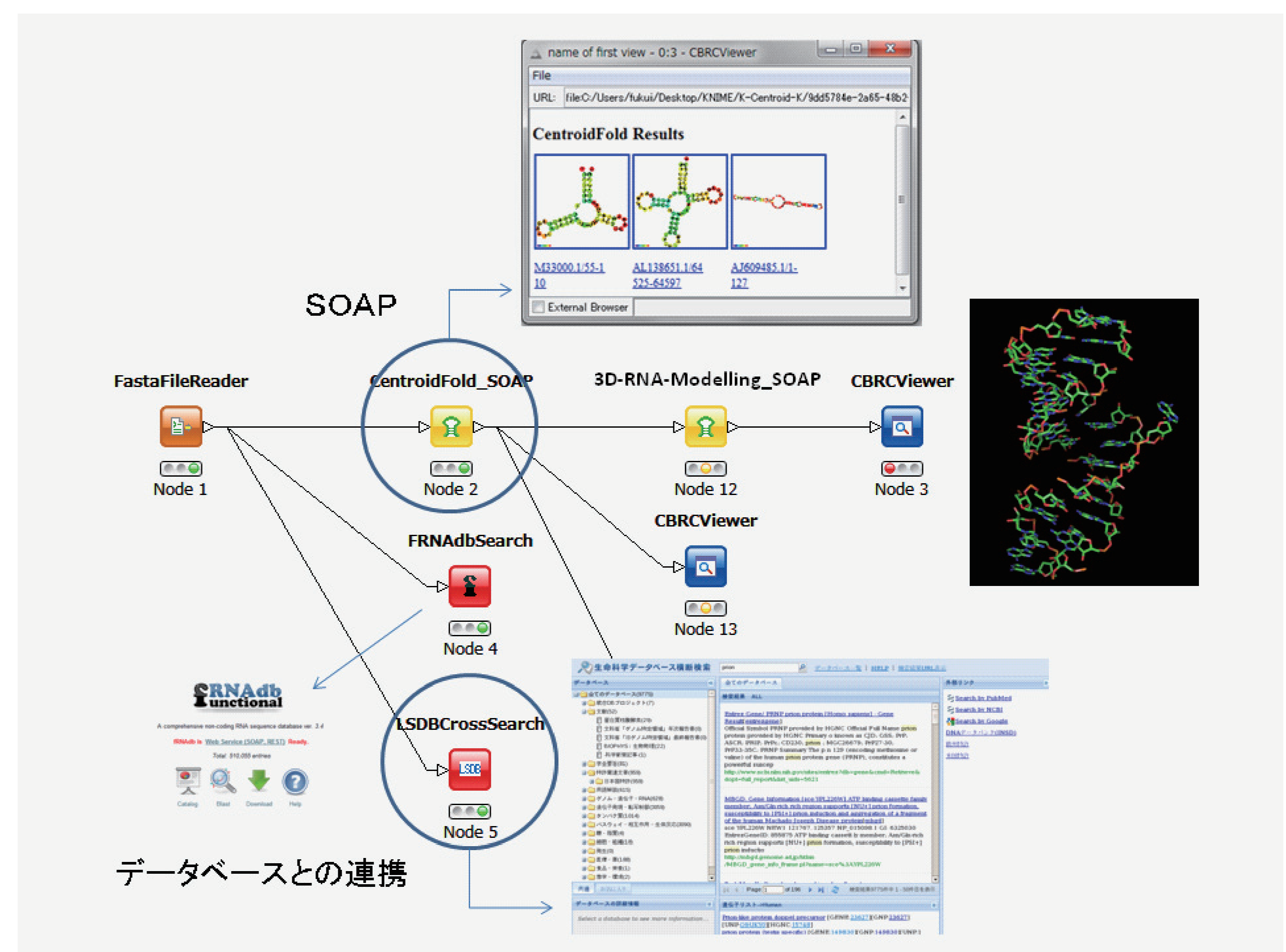
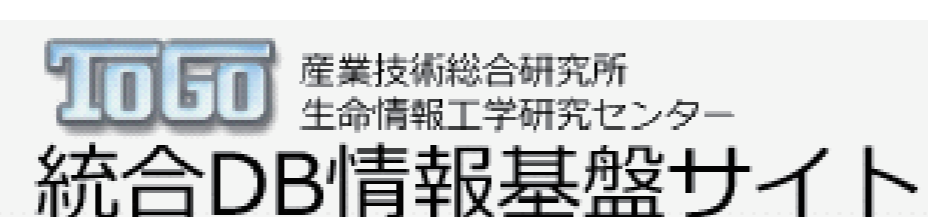


図3 アクティブ・ワークフロー

連絡先：分子機能計算チーム 福井一彦 k-fukui@aist.go.jp

<http://www.aist.go.jp/>