A-20

イネアノテーション計画データベース(RAP-DB)

田中剛、沼寿隆、坂井寛章、天野直己、伊藤剛

農業生物資源研究所　基盤研究領域　ゲノム情報研究ユニット

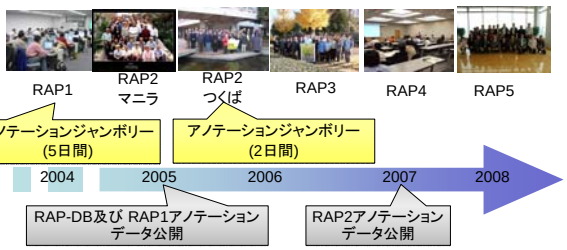
http://rapdb.dna.affrc.go.jp/

田中　剛: tstanaka@affrc.go.jp

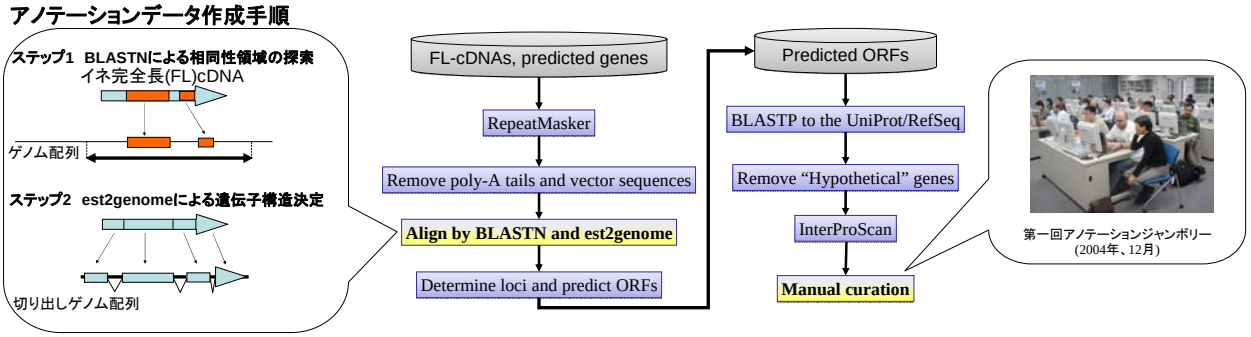
要旨

我々は、国際イネゲノム解読プロジェクト(IRGSP)の成果であるジャポニカイネ(*Oryza sativa* ssp. *japonica*)ゲノム配列決定をうけて、2004年からイネアノテーション計画(RAP)を遂行している。高精度のゲノムアノテーションデータを作成・公開することを本計画の目的として、データ作成とデータ公開用のデータベースの開発(RAP-DB)及び運営をこれまで行ってきた。遺伝子ごとの機能情報は精度向上のため、文献調査を手動で実行中である。また、遺伝子ファミリーデータを作成して公開したり、様々な研究機関によって作成されたリピート情報や、突然変異株情報などを随時RAP-DBから公開したりすることでイネ研究者がゲノム情報を基盤として、それぞれが必要な情報をRAP-DBから取得できるようにしている。現在は、イネゲノムの新バージョンが公開されたことを受け、新規アノテーションデータを作成しており、更にゲノムアラインメント・異種間転写産物マッピングを実行し種間ゲノム比較データ作成を行っている。これらのデータはα版としてRAP-DBから公開している。

RAPの歴史



アノテーションデータ作成手順



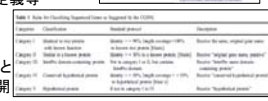
RAP2アノテーションデータ(公開中)

発現遺伝子座数	31,439
FLcDNA情報があるタンパク質コード遺伝子座	25,012
発現情報がある予測遺伝子	5,180
発現情報がある非タンパク質コード遺伝子座	1,247
発現情報無たんぱく質コード領域	22,022

カテゴリー	定義	ORF数
I	既知のイネ機能遺伝子と一致	664
II	既知機能遺伝子と相同	7,562
III	InterProドメインを持つ遺伝子	13,632
IV	機能未知遺伝子と相同	6,954
V	Hypothetical protein	1,380
合計		30,192

RAPは2004年より国際イネゲノム解読プロジェクト(IRGSP)や遺伝子命名グループ(CGSNL)と協調してアノテーションを推進している。

・IRGSP genomeに基づくアノテーション
・遺伝子座IDの振り方 (OsXXgXXXXXXX)
・遺伝子機能アノテーションの定義等



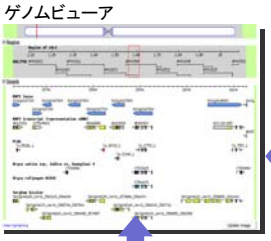
DDBJ/EMBL/GenBankやUniProtとも協力してアノテーションデータ公開

RAP-DB (http://rapdb.dna.affrc.go.jp/)

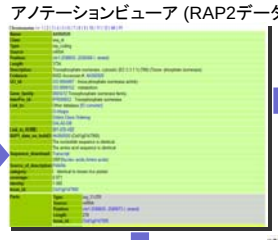
日本をイネ情報発信の中心にしたい！

トップページ

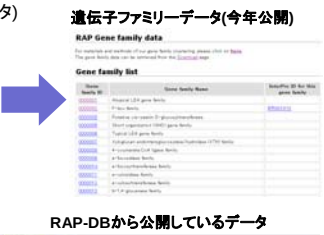
ゲノムビューア



アノテーションビューア (RAP2データ)



遺伝子ファミリーデータ(今年公開)

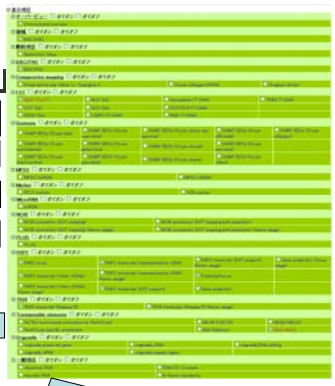


BLAST & キーワード検索, IDコンバータ

Annotation viewer (Comparative mapping)

他のDBへ

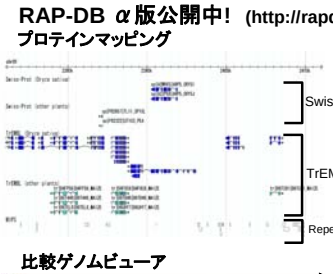
RAP-DBから公開しているデータ



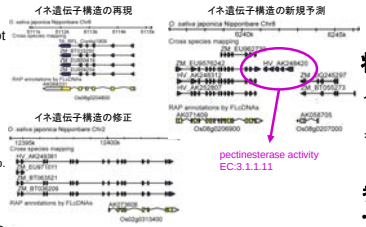
マーカー・オルガネラ情報なども更新

RAP-DB α版公開中! (http://rapdb.dna.affrc.go.jp/viewer/gbrowse/build5/)

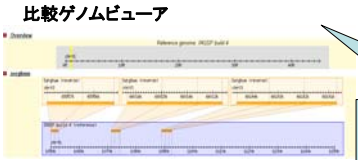
プロテインマッピング



異種マッピング



比較ゲノムビューア



ゲノムアノテーションのため、更にアノテーション情報の精度向上のために様々な情報を活用中(タンパク質データ・異種データ等)。

将来展望

イネゲノムの新バージョンが昨年度公開されたことから(IRGSP genome build 5)、新規アノテーションデータを作成中(RAP-DB α版より)。
＝＞ ゲノム情報は既に取得可。転写産物マッピング結果も閲覧が可能に。

参考文献

- Tanaka, T. et al., Nucleic Acids Res., 36(Database issue), D1028-33 (2008).
- Itoh, T., et al., Genome Research, 17, 175-183 (2007).

謝辞

本活動は農水受託(新農業展開プロジェクト、GIR-1001)の支援を受けて現在活動中です。イネアノテーション計画(RAP)に携わった全ての皆様に感謝いたします。