

機関名	課題項目	目標と達成状況																																																																																																																																																																																																														
		①委託事業の9月末時点の判断基準になる目標	②9月末時点の達成状況	③上記達成状況を踏まえたプロジェクト終了までの目標	④成果の概要																																																																																																																																																																																																											
DBCLS	(1)戦略立案・実行評価 (2)統合データベース開発 (3)共通基盤技術開発 (4)統合DBの開発・運用 (5)モデル生物・産業応用生物統合DBの開発・運用 (6)統合データベース支援 ①ポータル整備・運用、広報、普及啓発 ②データベースの受入と運用 (4)事業の総合的推進	・生命科学系データベースカタログ登録DB数:850件 ・生命科学系データベース横断検索対象DB数:250件 ・データベース受入れ数:38件(年度内では45件)	・生命科学系データベースカタログ登録DB数:912件(国内620件、国外292件) ・生命科学系データベース横断検索対象DB数:270件(国内150件、国外120件) ・データベース受入れ数:38件	・DBカタログ情報の経産省MEDALSとの共通化 ・横断検索IFの改良とデータ間のリンク情報による統合検索プロトタイプ構築 ・データベース受入れの最終目標45件を実現する。	・一意見訪問者 48,000名/月 ・ページビュー 1,000,000 ・ページ/月 (DBCLS提供サービス分)																																																																																																																																																																																																											
					<table border="1"> <caption>DBCLS成果のまとめ</caption> <thead> <tr> <th>分類</th> <th>項目</th> <th>概要</th> <th>利用状況(月間)</th> </tr> </thead> <tbody> <tr> <td>総合ポータルサイト</td> <td>統合DBプロジェクト構築の公開サイト、DBCLS各プロジェクト参加機関の公開サービス網のリンク、提供サービス7、12088月間開始、右はDBCLS分</td> <td></td> <td>968,768</td> </tr> <tr> <td rowspan="4">ポータル</td> <td>生命科学系</td> <td>DBサイトのトップページをサムネイルで表示、DB検索状況をもモニターして表示、利用</td> <td>300</td> </tr> <tr> <td>DBカタログ</td> <td>生命科学系DBカタログの公開、検索機能の強化、12088月間開始</td> <td>300</td> </tr> <tr> <td>生命科学系</td> <td>ゲノム、ホストゲノム関連を中心に30年代以降の関連の生命科学系の主要なプロジェクトの目録情報と検索、更新の強化、12088月間開始</td> <td>19</td> </tr> <tr> <td>主要ページ</td> <td>生命科学系DBカタログの目録情報と検索、更新の強化、12088月間開始</td> <td>30</td> </tr> <tr> <td rowspan="2">生物アイコン</td> <td>生命科学系DBカタログの目録情報と検索、更新の強化、12088月間開始</td> <td></td> <td>871</td> </tr> <tr> <td>生物アイコン</td> <td>生命科学系DBカタログの目録情報と検索、更新の強化、12088月間開始</td> <td>2,149</td> </tr> <tr> <td rowspan="2">新着論文レビュー</td> <td>生命科学系DBカタログの目録情報と検索、更新の強化、12088月間開始</td> <td></td> <td>3,845</td> </tr> <tr> <td>新着論文レビュー</td> <td>生命科学系DBカタログの目録情報と検索、更新の強化、12088月間開始</td> <td>42,815</td> </tr> <tr> <td rowspan="2">生命科学系DB</td> <td>生命科学系DB</td> <td>生命科学系DBカタログの目録情報と検索、更新の強化、12088月間開始</td> <td>13,632</td> </tr> <tr> <td>断片検索</td> <td>断片検索機能の実装、対象DB数270、12088月間開始</td> <td>4,987,877</td> </tr> <tr> <td rowspan="2">日本語コンテンツ公開</td> <td>日本語コンテンツ公開</td> <td>生命科学系DBカタログの目録情報と検索、更新の強化、12088月間開始</td> <td>4,984,409</td> </tr> <tr> <td>断片検索</td> <td>断片検索機能の実装、対象DB数270、12088月間開始</td> <td>19,918/10,048</td> </tr> <tr> <td rowspan="2">ToGoProt</td> <td>ToGoProt</td> <td>蛋白質データベース構築、対象DB数270、12088月間開始</td> <td>134</td> </tr> <tr> <td>断片検索</td> <td>断片検索機能の実装、対象DB数270、12088月間開始</td> <td>956</td> </tr> <tr> <td rowspan="2">ORFFL</td> <td>ORFFL</td> <td>ORFFLデータベース構築、対象DB数270、12088月間開始</td> <td>449</td> </tr> <tr> <td>断片検索</td> <td>断片検索機能の実装、対象DB数270、12088月間開始</td> <td>10,239</td> </tr> <tr> <td rowspan="2">Alfa</td> <td>Alfa</td> <td>Alfaデータベース構築、対象DB数270、12088月間開始</td> <td>1,235</td> </tr> <tr> <td>断片検索</td> <td>断片検索機能の実装、対象DB数270、12088月間開始</td> <td>6,659</td> </tr> <tr> <td rowspan="2">iMaYes</td> <td>iMaYes</td> <td>iMaYesデータベース構築、対象DB数270、12088月間開始</td> <td>751</td> </tr> <tr> <td>断片検索</td> <td>断片検索機能の実装、対象DB数270、12088月間開始</td> <td>11,772</td> </tr> <tr> <td rowspan="2">DNA データベース</td> <td>DNA データベース</td> <td>DNA データベース構築、対象DB数270、12088月間開始</td> <td>1,564</td> </tr> <tr> <td>断片検索</td> <td>断片検索機能の実装、対象DB数270、12088月間開始</td> <td>7,015</td> </tr> <tr> <td rowspan="2">遺伝子発現データ統合</td> <td>遺伝子発現データ統合</td> <td>遺伝子発現データベース構築、対象DB数270、12088月間開始</td> <td>244,770</td> </tr> <tr> <td>断片検索</td> <td>断片検索機能の実装、対象DB数270、12088月間開始</td> <td>1,564</td> </tr> <tr> <td rowspan="2">ToGoExp(RefEx)</td> <td>ToGoExp(RefEx)</td> <td>ToGoExp(RefEx)データベース構築、対象DB数270、12088月間開始</td> <td>159</td> </tr> <tr> <td>断片検索</td> <td>断片検索機能の実装、対象DB数270、12088月間開始</td> <td>1,711</td> </tr> <tr> <td rowspan="2">生命科学系</td> <td>生命科学系</td> <td>生命科学系データベース構築、対象DB数270、12088月間開始</td> <td>831</td> </tr> <tr> <td>断片検索</td> <td>断片検索機能の実装、対象DB数270、12088月間開始</td> <td>4,704</td> </tr> <tr> <td rowspan="2">BioPep-3D</td> <td>BioPep-3D</td> <td>BioPep-3Dデータベース構築、対象DB数270、12088月間開始</td> <td>4,893</td> </tr> <tr> <td>断片検索</td> <td>断片検索機能の実装、対象DB数270、12088月間開始</td> <td>248,853</td> </tr> <tr> <td rowspan="2">アノテーション</td> <td>アノテーション</td> <td>アノテーションデータベース構築、対象DB数270、12088月間開始</td> <td>368,183</td> </tr> <tr> <td>断片検索</td> <td>断片検索機能の実装、対象DB数270、12088月間開始</td> <td>1,977,000(※1)</td> </tr> <tr> <td rowspan="2">Wired-Marker</td> <td>Wired-Marker</td> <td>Wired-Markerデータベース構築、対象DB数270、12088月間開始</td> <td>196</td> </tr> <tr> <td>断片検索</td> <td>断片検索機能の実装、対象DB数270、12088月間開始</td> <td>13,532</td> </tr> <tr> <td rowspan="2">MGAP</td> <td>MGAP</td> <td>MGAPデータベース構築、対象DB数270、12088月間開始</td> <td>238</td> </tr> <tr> <td>断片検索</td> <td>断片検索機能の実装、対象DB数270、12088月間開始</td> <td>15,447</td> </tr> <tr> <td rowspan="2">DBCLS Galaxy</td> <td>DBCLS Galaxy</td> <td>DBCLS Galaxyデータベース構築、対象DB数270、12088月間開始</td> <td>1,302</td> </tr> <tr> <td>断片検索</td> <td>断片検索機能の実装、対象DB数270、12088月間開始</td> <td>110,165</td> </tr> <tr> <td rowspan="2">ToGoDB</td> <td>ToGoDB</td> <td>ToGoDBデータベース構築、対象DB数270、12088月間開始</td> <td>916</td> </tr> <tr> <td>断片検索</td> <td>断片検索機能の実装、対象DB数270、12088月間開始</td> <td>399,873</td> </tr> <tr> <td rowspan="2">ToGoWS</td> <td>ToGoWS</td> <td>ToGoWSデータベース構築、対象DB数270、12088月間開始</td> <td>387</td> </tr> <tr> <td>断片検索</td> <td>断片検索機能の実装、対象DB数270、12088月間開始</td> <td>24,417</td> </tr> <tr> <td rowspan="2">OpenID認証システム</td> <td>OpenID認証システム</td> <td>OpenID認証システム構築、対象DB数270、12088月間開始</td> <td>1,313</td> </tr> <tr> <td>断片検索</td> <td>断片検索機能の実装、対象DB数270、12088月間開始</td> <td>11,712</td> </tr> <tr> <td rowspan="2">BioHackathon</td> <td>BioHackathon</td> <td>BioHackathonデータベース構築、対象DB数270、12088月間開始</td> <td>993</td> </tr> <tr> <td>断片検索</td> <td>断片検索機能の実装、対象DB数270、12088月間開始</td> <td>10,549</td> </tr> <tr> <td rowspan="2">録音の構築と公開</td> <td>録音の構築と公開</td> <td>録音の構築と公開データベース構築、対象DB数270、12088月間開始</td> <td>916</td> </tr> <tr> <td>断片検索</td> <td>断片検索機能の実装、対象DB数270、12088月間開始</td> <td>6,008</td> </tr> <tr> <td rowspan="2">ToGoDoc</td> <td>ToGoDoc</td> <td>ToGoDocデータベース構築、対象DB数270、12088月間開始</td> <td>4,006</td> </tr> <tr> <td>断片検索</td> <td>断片検索機能の実装、対象DB数270、12088月間開始</td> <td>1,125</td> </tr> <tr> <td rowspan="2">ToGoTV</td> <td>ToGoTV</td> <td>ToGoTVデータベース構築、対象DB数270、12088月間開始</td> <td>1,125</td> </tr> <tr> <td>断片検索</td> <td>断片検索機能の実装、対象DB数270、12088月間開始</td> <td>1,125</td> </tr> <tr> <td rowspan="2">講習会(MoDB)</td> <td>講習会(MoDB)</td> <td>講習会(MoDB)データベース構築、対象DB数270、12088月間開始</td> <td>225</td> </tr> <tr> <td>断片検索</td> <td>断片検索機能の実装、対象DB数270、12088月間開始</td> <td>4,125</td> </tr> <tr> <td rowspan="2">外部評価</td> <td>外部評価</td> <td>外部評価データベース構築、対象DB数270、12088月間開始</td> <td>1,125</td> </tr> <tr> <td>断片検索</td> <td>断片検索機能の実装、対象DB数270、12088月間開始</td> <td>1,125</td> </tr> </tbody> </table>	分類	項目	概要	利用状況(月間)	総合ポータルサイト	統合DBプロジェクト構築の公開サイト、DBCLS各プロジェクト参加機関の公開サービス網のリンク、提供サービス7、12088月間開始、右はDBCLS分		968,768	ポータル	生命科学系	DBサイトのトップページをサムネイルで表示、DB検索状況をもモニターして表示、利用	300	DBカタログ	生命科学系DBカタログの公開、検索機能の強化、12088月間開始	300	生命科学系	ゲノム、ホストゲノム関連を中心に30年代以降の関連の生命科学系の主要なプロジェクトの目録情報と検索、更新の強化、12088月間開始	19	主要ページ	生命科学系DBカタログの目録情報と検索、更新の強化、12088月間開始	30	生物アイコン	生命科学系DBカタログの目録情報と検索、更新の強化、12088月間開始		871	生物アイコン	生命科学系DBカタログの目録情報と検索、更新の強化、12088月間開始	2,149	新着論文レビュー	生命科学系DBカタログの目録情報と検索、更新の強化、12088月間開始		3,845	新着論文レビュー	生命科学系DBカタログの目録情報と検索、更新の強化、12088月間開始	42,815	生命科学系DB	生命科学系DB	生命科学系DBカタログの目録情報と検索、更新の強化、12088月間開始	13,632	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	4,987,877	日本語コンテンツ公開	日本語コンテンツ公開	生命科学系DBカタログの目録情報と検索、更新の強化、12088月間開始	4,984,409	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	19,918/10,048	ToGoProt	ToGoProt	蛋白質データベース構築、対象DB数270、12088月間開始	134	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	956	ORFFL	ORFFL	ORFFLデータベース構築、対象DB数270、12088月間開始	449	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	10,239	Alfa	Alfa	Alfaデータベース構築、対象DB数270、12088月間開始	1,235	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	6,659	iMaYes	iMaYes	iMaYesデータベース構築、対象DB数270、12088月間開始	751	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	11,772	DNA データベース	DNA データベース	DNA データベース構築、対象DB数270、12088月間開始	1,564	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	7,015	遺伝子発現データ統合	遺伝子発現データ統合	遺伝子発現データベース構築、対象DB数270、12088月間開始	244,770	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	1,564	ToGoExp(RefEx)	ToGoExp(RefEx)	ToGoExp(RefEx)データベース構築、対象DB数270、12088月間開始	159	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	1,711	生命科学系	生命科学系	生命科学系データベース構築、対象DB数270、12088月間開始	831	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	4,704	BioPep-3D	BioPep-3D	BioPep-3Dデータベース構築、対象DB数270、12088月間開始	4,893	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	248,853	アノテーション	アノテーション	アノテーションデータベース構築、対象DB数270、12088月間開始	368,183	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	1,977,000(※1)	Wired-Marker	Wired-Marker	Wired-Markerデータベース構築、対象DB数270、12088月間開始	196	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	13,532	MGAP	MGAP	MGAPデータベース構築、対象DB数270、12088月間開始	238	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	15,447	DBCLS Galaxy	DBCLS Galaxy	DBCLS Galaxyデータベース構築、対象DB数270、12088月間開始	1,302	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	110,165	ToGoDB	ToGoDB	ToGoDBデータベース構築、対象DB数270、12088月間開始	916	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	399,873	ToGoWS	ToGoWS	ToGoWSデータベース構築、対象DB数270、12088月間開始	387	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	24,417	OpenID認証システム	OpenID認証システム	OpenID認証システム構築、対象DB数270、12088月間開始	1,313	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	11,712	BioHackathon	BioHackathon	BioHackathonデータベース構築、対象DB数270、12088月間開始	993	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	10,549	録音の構築と公開	録音の構築と公開	録音の構築と公開データベース構築、対象DB数270、12088月間開始	916	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	6,008	ToGoDoc	ToGoDoc	ToGoDocデータベース構築、対象DB数270、12088月間開始	4,006	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	1,125	ToGoTV	ToGoTV	ToGoTVデータベース構築、対象DB数270、12088月間開始	1,125	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	1,125	講習会(MoDB)	講習会(MoDB)	講習会(MoDB)データベース構築、対象DB数270、12088月間開始	225	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	4,125	外部評価	外部評価	外部評価データベース構築、対象DB数270、12088月間開始	1,125	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	1,125
分類	項目	概要	利用状況(月間)																																																																																																																																																																																																													
総合ポータルサイト	統合DBプロジェクト構築の公開サイト、DBCLS各プロジェクト参加機関の公開サービス網のリンク、提供サービス7、12088月間開始、右はDBCLS分		968,768																																																																																																																																																																																																													
ポータル	生命科学系	DBサイトのトップページをサムネイルで表示、DB検索状況をもモニターして表示、利用	300																																																																																																																																																																																																													
	DBカタログ	生命科学系DBカタログの公開、検索機能の強化、12088月間開始	300																																																																																																																																																																																																													
	生命科学系	ゲノム、ホストゲノム関連を中心に30年代以降の関連の生命科学系の主要なプロジェクトの目録情報と検索、更新の強化、12088月間開始	19																																																																																																																																																																																																													
	主要ページ	生命科学系DBカタログの目録情報と検索、更新の強化、12088月間開始	30																																																																																																																																																																																																													
生物アイコン	生命科学系DBカタログの目録情報と検索、更新の強化、12088月間開始		871																																																																																																																																																																																																													
	生物アイコン	生命科学系DBカタログの目録情報と検索、更新の強化、12088月間開始	2,149																																																																																																																																																																																																													
新着論文レビュー	生命科学系DBカタログの目録情報と検索、更新の強化、12088月間開始		3,845																																																																																																																																																																																																													
	新着論文レビュー	生命科学系DBカタログの目録情報と検索、更新の強化、12088月間開始	42,815																																																																																																																																																																																																													
生命科学系DB	生命科学系DB	生命科学系DBカタログの目録情報と検索、更新の強化、12088月間開始	13,632																																																																																																																																																																																																													
	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	4,987,877																																																																																																																																																																																																													
日本語コンテンツ公開	日本語コンテンツ公開	生命科学系DBカタログの目録情報と検索、更新の強化、12088月間開始	4,984,409																																																																																																																																																																																																													
	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	19,918/10,048																																																																																																																																																																																																													
ToGoProt	ToGoProt	蛋白質データベース構築、対象DB数270、12088月間開始	134																																																																																																																																																																																																													
	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	956																																																																																																																																																																																																													
ORFFL	ORFFL	ORFFLデータベース構築、対象DB数270、12088月間開始	449																																																																																																																																																																																																													
	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	10,239																																																																																																																																																																																																													
Alfa	Alfa	Alfaデータベース構築、対象DB数270、12088月間開始	1,235																																																																																																																																																																																																													
	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	6,659																																																																																																																																																																																																													
iMaYes	iMaYes	iMaYesデータベース構築、対象DB数270、12088月間開始	751																																																																																																																																																																																																													
	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	11,772																																																																																																																																																																																																													
DNA データベース	DNA データベース	DNA データベース構築、対象DB数270、12088月間開始	1,564																																																																																																																																																																																																													
	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	7,015																																																																																																																																																																																																													
遺伝子発現データ統合	遺伝子発現データ統合	遺伝子発現データベース構築、対象DB数270、12088月間開始	244,770																																																																																																																																																																																																													
	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	1,564																																																																																																																																																																																																													
ToGoExp(RefEx)	ToGoExp(RefEx)	ToGoExp(RefEx)データベース構築、対象DB数270、12088月間開始	159																																																																																																																																																																																																													
	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	1,711																																																																																																																																																																																																													
生命科学系	生命科学系	生命科学系データベース構築、対象DB数270、12088月間開始	831																																																																																																																																																																																																													
	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	4,704																																																																																																																																																																																																													
BioPep-3D	BioPep-3D	BioPep-3Dデータベース構築、対象DB数270、12088月間開始	4,893																																																																																																																																																																																																													
	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	248,853																																																																																																																																																																																																													
アノテーション	アノテーション	アノテーションデータベース構築、対象DB数270、12088月間開始	368,183																																																																																																																																																																																																													
	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	1,977,000(※1)																																																																																																																																																																																																													
Wired-Marker	Wired-Marker	Wired-Markerデータベース構築、対象DB数270、12088月間開始	196																																																																																																																																																																																																													
	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	13,532																																																																																																																																																																																																													
MGAP	MGAP	MGAPデータベース構築、対象DB数270、12088月間開始	238																																																																																																																																																																																																													
	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	15,447																																																																																																																																																																																																													
DBCLS Galaxy	DBCLS Galaxy	DBCLS Galaxyデータベース構築、対象DB数270、12088月間開始	1,302																																																																																																																																																																																																													
	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	110,165																																																																																																																																																																																																													
ToGoDB	ToGoDB	ToGoDBデータベース構築、対象DB数270、12088月間開始	916																																																																																																																																																																																																													
	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	399,873																																																																																																																																																																																																													
ToGoWS	ToGoWS	ToGoWSデータベース構築、対象DB数270、12088月間開始	387																																																																																																																																																																																																													
	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	24,417																																																																																																																																																																																																													
OpenID認証システム	OpenID認証システム	OpenID認証システム構築、対象DB数270、12088月間開始	1,313																																																																																																																																																																																																													
	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	11,712																																																																																																																																																																																																													
BioHackathon	BioHackathon	BioHackathonデータベース構築、対象DB数270、12088月間開始	993																																																																																																																																																																																																													
	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	10,549																																																																																																																																																																																																													
録音の構築と公開	録音の構築と公開	録音の構築と公開データベース構築、対象DB数270、12088月間開始	916																																																																																																																																																																																																													
	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	6,008																																																																																																																																																																																																													
ToGoDoc	ToGoDoc	ToGoDocデータベース構築、対象DB数270、12088月間開始	4,006																																																																																																																																																																																																													
	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	1,125																																																																																																																																																																																																													
ToGoTV	ToGoTV	ToGoTVデータベース構築、対象DB数270、12088月間開始	1,125																																																																																																																																																																																																													
	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	1,125																																																																																																																																																																																																													
講習会(MoDB)	講習会(MoDB)	講習会(MoDB)データベース構築、対象DB数270、12088月間開始	225																																																																																																																																																																																																													
	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	4,125																																																																																																																																																																																																													
外部評価	外部評価	外部評価データベース構築、対象DB数270、12088月間開始	1,125																																																																																																																																																																																																													
	断片検索	断片検索機能の実装、対象DB数270、12088月間開始	1,125																																																																																																																																																																																																													
JST	1)継続運用 ①意見集約システムの運用 ②事業サイト等の運用 ③国際標準メタデータエレメントサイトの運用 2)統合データベース運用の委託	・意見集約システムを継続して運用する。 Wikiへの参加を誘発するために、(例えばNCBIなど諸外国の)新規DBデータを5DB追加する。 ・事業サイト等を継続して運用する。 ・メタデータエレメントサイト(MDeR)を継続して運用し、併せて中核機関のアーカイブサイトにサンプルデータを提供する。	・以下5つのデータベースをWINGprolに追加した。 1. Probe Database(8月26日公開) 2. Epigenomics(9月21日公開) 3. Peptidome(9月14日公開) 4. PRIDE(PRoteomics IDEntifications database)(9月21日公開) 5. MassBank(9月15日公開) ・事業サイトの継続運用を実施した。 ・MDeRの継続運用を実施し、中核機関のアーカイブサイトに対しサンプルデータ提供を行った。	・意見集約システムを継続運用する。 ・事業サイト等を継続運用する。事業サイトは最終的に「ライフサイエンスの広場」に静的にリンクし保存する。 ・メタデータエレメントサイトを継続運用する。	WINGprolに追加した新規データベースは1〜3がNCBI、4がEBI、5が慶應義塾大学のデータベース。2010年10月19日現在までのアクセス数は各々98件、73件、116件、81件、121件。 2010年4月から9月末までのアクセス件数は以下の通り。 ・WINGpro :17971件(前年同期間:21584件) ・事業サイト:10587件(前年同期間:10318件) ・MDeR :1772件(前年同期間:1419件)																																																																																																																																																																																																											
かずさDNA研	①高度情報集積データベースの運用と改良 ②ゲノムアノテーション情報の蓄積と高度化	・「高度情報集積データベース」については、Gene Indexing型アノテーション情報を、複合体への注釈を含め18万件以上の蓄積を完了する。 ・植物と関連微生物のコミュニティへの周知を徹底するとともに基盤データベースと閲覧環境の改善をすすめ、100名規模のアクティブな利用者を確保することを目標とする。 ・「BioMartによる植物と関連微生物ゲノム情報の統合」については、イネ、シロイヌナズナをはじめとした解析情報の整備された主要植物8種ならびに植物関連微生物16種類のゲノム情報を追加し、統合利用環境の整備を完了し広く利用に供する。	・Gene Indexing型アノテーション180003件(9/24時点)達成。 ・ナビゲーション(研究者SNS)の登録利用者115名達成。土壌細菌コミュニティでの新規ゲノムアノテーションジャンボリーでの利用。 ・BioMart植物関連微生物18種類、主要植物8種類追加。統合データベース講習会や学会での広報活動。	・Gene Indexing型、Complex Indexing型アノテーションを合計して19万件以上の蓄積。 ・かずさナビゲーション(研究者SNS)登録利用者120名。 ・BioMart植物関連微生物5種類、主要植物4種類追加。	1)高度情報集積データベースの運用と改良 平成22年度の新しいとくみとして、KazusaAnnotationを利用して、植物関連土壌細菌のゲノムアノテーションジャンボリーを行った。アノテーションジャンボリーでは、公開前のゲノム情報を扱う。そのため、ジャンボリー参加者のみが情報を共有できる仕組みをKazusaAnnotationに追加した。また、複数人数で一貫性のあるゲノムアノテーションを行うための標準作業手順を開発しKazusaWiki(http://wiki.kazusa.or.jp)によりマニュアルや資料の作成を共有した。かずさの提供するシステム組み合わせることにより、研究コミュニティによるマニュアルアノテーションを成功させた。 2)ゲノムアノテーション情報の蓄積と高度化 平成22年度も論文中出现する遺伝子・タンパク質記載のインデックス情報(Gene Indexing型アノテーション)をKazusaAnnotationを利用し蓄積した。 3)植物と関連微生物ゲノム情報の統合 生物種をまたいだ横断的なゲノム情報の取得や、理化学研究所など他機関の植物オームクス情報との連携のため、KazusaMart(http://mart.kazusa.or.jp)を構築した。																																																																																																																																																																																																											
中核機関 産総研CBRC	①ワークフロー技術を用いた統合DB環境の構築	目標値と内容:ワークフロー技術を用いた統合DB環境構築とし、 ・3つのタンパク質に関連したWEBワークフローを公開 ・CBRCが独自に開発した解析ツールを用い、5つのアクティブワークフローを公開 ・アクティブワークフロー上で2つ以上の研究機関(DBCLS等)との連携を行うノードの開発	・タンパク質関連の3つのWEBワークフローを公開 ・SOAP通信等を用いた8つのアクティブワークフローを公開 ・DBCLS、DDBj、PDBjのサービスと連携したノードの開発	・これまで開発したノードと既存のノードを繋げたケーススタディの公開 ・LINUXやMACにも対応したノードの開発	WEBワークフロー (http://togo.cbrc.jp/web.html) ・Protein Annotation Workflow (2008年 12月) ・Comparative Protein Information Workflow (2009年3月) ・Protein Modelling Workflow (2009年 11月) Activeワークフロー (http://togo.cbrc.jp/active.html) ・WolfPSORT Active Workflow (2010年8月) ・Last Active Workflow (2010年8月) ・Blast Active Workflow (2010年8月) ・Mafft Active Workflow (2010年8月) ・Modelling Active Workflow (2010年7月) ・CentroidFold Active Workflow (2010年6月) ・POODLE Active Workflow (2010年5月) ・ASIAN Active Workflow (2010年 1月) プラットフォーム(ケーススタディ) ・RNA/ Protein Structure Prediction Workflow (2010年10月) ・PhylogeneticTree Active Workflow (2010年9月) アクセス数:約160/月 月間転送量:126MB(7月)200MB(8.9月)760MB(10月)																																																																																																																																																																																																											
奈良先端	(1)専門用語獲得技術 ①専門用語辞書システムの開発 ②専門用語解析技術の開発 (2)専門用語抽出ツールの設計と開発 ①専門用語辞書拡張支援ツールの設計と開発	・専門用語辞書システムを、属性情報等の追加に柔軟に対応するため、MongoDBでの実装を完成させる。 ・専門用語の内部構造解析済みデータを2500語以上に拡大する。 ・専門用語辞書システムとシソーラスの階層構造表示システムを統合する。	・専門用語辞書システムの属性情報追加機能は設計(仕様)を完了。MongoDBへの実装は進行中。 ・専門用語の内部構造解析済みデータの蓄積:9月末時点で2656語。(10月28日現在で、3454語) ・専門用語辞書システムとシソーラスの階層構造表示システムの統合:シソーラス階層表示システムから専門用語辞書システム(Cradle-LSD)へのリンクは完成。	・専門用語辞書システムのMongoDBへの移植はプロジェクト終了までには達成。 ・専門用語内部構造解析済みデータは、今後も若干作業を予定しており、少なくとも4000語規模に拡大する予定。 ・専門用語辞書システムからシソーラス階層構造表示システムへのリンクを実装し、両システムの統合を完成させる。	ユーザー評価サイトで、以下の2種類3つのシステムを公開した。公開日はいずれも10月25日。 ・ライフサイエンス辞書検索システム: Cradle-LSD ライフサイエンス辞書を対象に、用語そのものだけでなく、部分文字列、品詞、シソーラスコードなど様々な情報によって用語検索を行う。また、用語の内部構造の表示や内部構造アノテーションを行う機能をもつ。 ・バイオ医療専門用語と類似文脈の検索 ・PNE と LSD に対するバイオ医療専門用語の検索 指定した用語のPNE(雑誌/蛋白質・核酸・酵素)内での出現を検索し、前後文脈とともに表示する。また、その用語のMeSHシソーラス内での位置を表示する。各シソーラスコードには対応する専門用語(日本語)が表記され、Cradle-LSDにリンクされている。 ・PNE に対する類似文脈の検索 専門用語とその出現文脈を入力すると、類似した文脈をPNEより検索し、それによって用語と類似の専門用語を検索するシステム																																																																																																																																																																																																											
九大	①多型知識表現技術の開発	・9月末には2個以上のGWASのQC結果をJAGQCで公開する	・4個のGWASのQC結果をJAGQCで公開した。	・JAGQCパイプラインの拡充。 ・未取得GWASデータの取得。	データ請求件数:7(うち3件はデータ未取得) データ解析完了件数:4 解析項目:各取得データ当たり6項目(結果表示16細目) サービス公開日:2010年9月30日(アクセスはデータ提供者に制限) 同10月6日~ 公開 アクセス件数:未集計(公開日数が短いため)																																																																																																																																																																																																											

機関名	課題項目	目標と達成状況																																																																																			
		①委託事業の9月末時点の判断基準になる目標	②9月末時点の達成状況	③上記達成状況を踏まえたプロジェクト終了までの目標	④成果の概要																																																																																
東大	①バイオ DB サーバー構築演習 ②独自のサーバー構築演習	・大規模計算のためのクラスター利用技術と独自のバイオDBサーバーの構築技術を指導し、次世代シーケンシング情報のデータベースを構築できる人材を育成する。ウェブで公開中の演習ノートも改善する。本年度の受講者数は東大新領域4名、4名の受講者が次世代シーケンサー解析技術の基礎を習得することを9月末時点での目標値とする。	・カリキュラムとして設定した1回生、2回生、3回生の実習を行った。 ・tRNAデータベースの更新と機能拡充についての最新論文"tRNADB-CE 2011: tRNA gene database curated manually by experts"が、Nucl Acid Resに採択されたので、2011年1月号に掲載される。 ・松本大学にて、7月26日に「健康への貢献遺伝子探索」の事前実習を行った。受講者は、松本大学の学部生と諏訪青陵高校の学生の合計で25名、ならびに教員が2名。 ・東邦大学理学部の実習で、「持続可能型社会への貢献遺伝子探索」のテキストが採用された。 ・本学のオープンキャンパス(8月8日と22日)に、統合DB整備事業とその成果を紹介した。 ・虎姫高校の40名を対象にして、9月11日に「病気に関する遺伝子探索」の実習授業を行った。	・演習ノートをさらに整理・改善し、ウェブ上の自習用教材として活用できるように演習資料を整備して公開することをプロジェクト終了までの目標とする。	演習内容を演習ノートとしてウェブ上に公開した。今年度の演習ノートのアクセス統計は以下の通りだった。(Sites: いくつかのIPアドレスからリクエストを受けたかを示す。KBytes: 送信したデータ量。Visits: サイトを訪問したユーザー数。Pages: 送信したページ数。画像などのファイルは含まない。Files: 送信したファイル数。) <table border="1"> <thead> <tr> <th rowspan="2">Month</th> <th colspan="4">Daily Avg</th> <th colspan="4">Monthly Totals</th> </tr> <tr> <th>Files</th> <th>Pages</th> <th>Visits</th> <th>Sites</th> <th>KBytes</th> <th>Visits</th> <th>Pages</th> <th>Files</th> </tr> </thead> <tbody> <tr> <td>Oct</td> <td>347</td> <td>321</td> <td>250</td> <td>300</td> <td>5829</td> <td>3764</td> <td>4520</td> <td>5217</td> </tr> <tr> <td>Sep</td> <td>257</td> <td>237</td> <td>246</td> <td>630</td> <td>16944</td> <td>7380</td> <td>10125</td> <td>10730</td> </tr> <tr> <td>Aug</td> <td>1096</td> <td>1073</td> <td>315</td> <td>978</td> <td>40620</td> <td>9778</td> <td>32287</td> <td>34066</td> </tr> <tr> <td>Jul</td> <td>1782</td> <td>1891</td> <td>488</td> <td>1938</td> <td>779913</td> <td>15927</td> <td>58823</td> <td>55244</td> </tr> <tr> <td>Jun</td> <td>580</td> <td>1108</td> <td>420</td> <td>1432</td> <td>318922</td> <td>12628</td> <td>33258</td> <td>17428</td> </tr> <tr> <td>May</td> <td>1515</td> <td>1575</td> <td>483</td> <td>2230</td> <td>688661</td> <td>15002</td> <td>48825</td> <td>46888</td> </tr> <tr> <td>Apr</td> <td>1382</td> <td>1354</td> <td>344</td> <td>1538</td> <td>614311</td> <td>10210</td> <td>40644</td> <td>40863</td> </tr> </tbody> </table> (10月のアクセス統計は10月15日までのアクセスログから作成。)	Month	Daily Avg				Monthly Totals				Files	Pages	Visits	Sites	KBytes	Visits	Pages	Files	Oct	347	321	250	300	5829	3764	4520	5217	Sep	257	237	246	630	16944	7380	10125	10730	Aug	1096	1073	315	978	40620	9778	32287	34066	Jul	1782	1891	488	1938	779913	15927	58823	55244	Jun	580	1108	420	1432	318922	12628	33258	17428	May	1515	1575	483	2230	688661	15002	48825	46888	Apr	1382	1354	344	1538	614311	10210	40644	40863
Month	Daily Avg				Monthly Totals																																																																																
	Files	Pages	Visits	Sites	KBytes	Visits	Pages	Files																																																																													
Oct	347	321	250	300	5829	3764	4520	5217																																																																													
Sep	257	237	246	630	16944	7380	10125	10730																																																																													
Aug	1096	1073	315	978	40620	9778	32287	34066																																																																													
Jul	1782	1891	488	1938	779913	15927	58823	55244																																																																													
Jun	580	1108	420	1432	318922	12628	33258	17428																																																																													
May	1515	1575	483	2230	688661	15002	48825	46888																																																																													
Apr	1382	1354	344	1538	614311	10210	40644	40863																																																																													
長浜	①アノテータ・キュレータの教育 ②自己組織化マップによる大量ゲノム配列データへのアノテーション手法を用いた人材育成 ③シニア研究者のチームによる分散型の高度なキュレーションシステムの改良・運用	・1回生、2回生、3回生の実習に関しては、すでにカリキュラムとして設定しているため、完全に実施される。 ・tRNAデータベースの更新と機能拡充に関するキュレータ養成に関しては、すでに作業を完了しているが、この新バージョンに関する紹介論文がnucleic acid researchのデータベース特集号に掲載されることを目標としている。 ・他大学での事前実習については、9月までには1大学程度を想定しているが、それが達成できれば十分と考える。 ・中学生、高校生、高専生を対象にした講習会については、学内での開催を含めて3件程度が達成出来れば良いと考えている。	・カリキュラムとして設定した3回生の「次世代シーケンサーで解読した大量ゲノム情報のアノテーション実習」と、1回生を対象にした「バイオ技術の統合的データベース作成のための情報収集」を行う。	NARのDB特集号(2011)での発表が受理された。「tRNADB-CE 2011: tRNA gene database curated manually by experts」の場合、入力コード数が154,455レコード、データ量は39Mbで、アクセス数は年間で180万件程度である。 学生実習で作成した、「健康への貢献遺伝子データベース」と「持続可能型社会への貢献遺伝子データベース」とその実習テキストは、学内ホームページだけでなく、統合ホームページからも公開している。																																																																																	
お茶大	①DB高度利用者の養成	・9月時点で社会人2名を含む7名以上への講習を行う ・3名以上が継続して統合データベース事業に従事する	・講習の実施 毎週月曜日(前期): ネットワークを通じたライフサイエンスデータの利用 8月集中 ネットワーク基礎、現代遺伝学、データベース基礎、ネットワークを通じたライフサイエンスデータの利用、データ解析の基礎の講義を実施 社会人8名を含む16名が受講、そのうち4名の社会人は部分受講。 ・統合データベースセンターで2名、リサーチアシスタントとしてテキスト処理事業に従事。 ・お茶の水女子大学で2名が統合データベースのサーバ管理に従事。 ・ネットワークを通じたライフサイエンスデータの利用、データ解析基礎、データマイニングは、昨年度よりSlideshareを通じて公開	・講習の継続実施(後期木曜日・データマイニング) ・受講生は現時点で目標人数を超えているので、継続した実施が重要 ・本事業への従事者も、当初目標を上回る人数のため、継続して従事できるよう見守る ・ネットワーク基礎の講義録をダウンロード可能なpdf形式にて公開(10月12日より公開中) ・データベース基礎の講義録の公開、現在準備中(11月2日公開予定) ・DDBJの中村保一先生をお呼びして講演会の実施(2011年1月13日(木))	・ネットワークを通じたライフサイエンスデータの利用、データ解析基礎、データマイニングの講習資料を昨年度より公開。現代遺伝学は参考資料を作成し、本年度より公開。 ・ネットワーク基礎の講義録をダウンロード可能なpdf形式にて10月12日より公開中。 ・データベース基礎の講義録については、現在準備中(11月8日公開予定) ・統合DB事業への従事者の輩出(テキスト処理2名、統合TV2名、サーバ管理5名)社会人及び主婦受講生の獲得(アステラス、カコメ、ジナリス、虎ノ門病院、かずさアノテータ等) ・ホームページへのアクセス数 <table border="1"> <thead> <tr> <th>期間</th> <th>訪問者数</th> <th>一意な訪問者数</th> <th>ページビュー</th> </tr> </thead> <tbody> <tr> <td>1-3月</td> <td>1,474</td> <td>952</td> <td>5397</td> </tr> <tr> <td>4-6月</td> <td>2,504</td> <td>1,821</td> <td>5,524</td> </tr> <tr> <td>7-9月</td> <td>3,179</td> <td>2,116</td> <td>7,068</td> </tr> </tbody> </table>	期間	訪問者数	一意な訪問者数	ページビュー	1-3月	1,474	952	5397	4-6月	2,504	1,821	5,524	7-9月	3,179	2,116	7,068																																																																
期間	訪問者数	一意な訪問者数	ページビュー																																																																																		
1-3月	1,474	952	5397																																																																																		
4-6月	2,504	1,821	5,524																																																																																		
7-9月	3,179	2,116	7,068																																																																																		
京大	(1)共通基盤技術開発 (2)統合データベース開発・運用 (3)プロジェクトの総合的推進	・ターゲット化合物を全対全で類似構造検索するためのインタフェースを開発。 ・類似構造検索パッケージの公開 ・化学反応ネットワーク予測システムの改良 ・LinkDB検索インタフェースの改良	・化合物データセット間の構造比較を全体で行うためのインタフェースSIMCOMP2を公開した。 ・類似構造検索ツール群をパッケージ化しGenomeNetのFTPサイトに公開した(10月)。 ・化学反応ネットワーク予測システムを改良し、類似反応経路検索システムを開発した(10月)。 ・医薬品データベースの構造検索結果でアライメント表示するように改良した。 ・医薬品データベースのお薬手帖で併用注意・併用禁忌の情報を表示するように改良した。	・化学構造検索、化学反応ネットワーク予測ツールともに、要素技術とそのインタフェースの開発は完了したので、化学情報解析ツールをまとめたインタフェースを作成するとともに、改良点を洗い出す。 ・統合データベース開発・運用では、開発してきた医薬品データベースの関連情報の整備を引き続き進めるとともに、各種データベースの更新は随時進める。	知識処理技術開発では、化合物データセット間の構造比較を全体で行うためのインタフェースSIMCOMP2を開発し、http://www.genome.jp/tools/simcomp2/から8月に公開した。また、医薬品データベースに対する化学構造検索システムを6月に公開し、9月には検索結果で構造アライメントを表示するように改良したものを公開した。これらの構造検索で用いているプログラム群をパッケージ化し、GenomeNetのFTPサイト(http://ftp.genome.jp/tools/)で10月に公開した。反応ネットワーク予測システムPathPredを開発し1月に公開し、反応の基質と生成物のペアを抽出したデータベースRPAIRとペアから抽出した反応パターンデータベースRCLASS(2010年10月29日版でそれぞれ12,333と2,302エントリー)を整備している。これらのデータベースを活用した類似反応経路検索システムを開発し9月末時点でテスト公開している。化合物解析システム関係のツールを集めたサイトのアクセス数は、2010年9月で3,016ノードから約25万アクセス(ロボット含む)であった。統合データベース開発・運用では、主に医薬品データベースの改良を進めてきた。6月には個人用のおくすり手帖を公開するとともに、各有効成分からターゲットや代謝酵素の情報に関してKEGGのPATHWAYなどにリンクを張った。7月にはJ-GLOBALへのリンクを追加し、9月にはおくすり手帖で併用注意・併用禁忌の情報を表示するようにした。2010年9月版で医療用医薬品12,239件、一般用医薬品12,042件登録されている。医薬品データベースへのアクセス数は2010年9月で18,380ノードから約88万アクセス(ロボット含む)であった。																																																																																
分担機関	医科歯科大 ①疾患・臨床医学データベースに特化した統合技術開発に関する研究(がん疾患を中心として) ②疾患データベースの高度化と実証的統合に関する研究(がん疾患を中心として) ③プロジェクトの総合的推進 東京医科歯科大 ・阪大 ①疾患・臨床医学データベースに特化した統合技術開発に関する研究(神経疾患について) ②疾患データベースの高度化と実証的統合に関する研究(神経疾患について)	・疾患を限定し、実用に耐えられる完成度をもった各要素データベース間の横断的検索システムの整備 ・分子レベルから臨床レベルの情報を検索し、症例データ及び分子・臨床レベルの情報を統合的に表示する機能の高度化 ・倫理規定ガイドラインにおける公開情報の階層化や利用者の類別に従ったアクセス権限などの検討 ・統合検索システムを介した神経疾患(パーキンソン病)データベースの検索機能の開発・改良作業を完了させる。 ・大阪大学の神経疾患データベースについて、9月時点で公開されている関連情報とのリンク付け、および関連機関への情報提供を完了させる。	・必要最小限の機能を備えたシステムの開発を終え、成果物として横断的統合検索用APIを公開予定 ・分子レベルから臨床レベルの縦断的統合検索システム(二次元三層マップ、遺伝子検索など)の整備を終え、最終的には成果物としてプログラム等を公開予定 ・倫理規定ガイドラインにデータ種類別アクセスレベル等の内容を追記	・横断的・縦断的統合検索用のAPIやプログラム等を成果物として公開 ・倫理規定ガイドラインの提案 ・関連情報へのリンク付けなど、広報活動の推進 ・論文掲載	2010年9月30日に完全公開を開始。アクセス数については公開後間もないため、今回は成果物の対象としない。縦断的統合はiODDとGeMDBJのデータより、公開版(140症例)、登録版(447症例)が対象となっている。																																																																																
東大G ・東大医学 ・東大病院 ・東海大 ・日立	(1)標準SNP DBの構築 (2)GWAS DBの構築 (3)GWAS 第1ステージDBの構築 (4)リシーケンスによる臨床情報・ゲノム情報DBの構築 ①リシーケンスDBへの外部DBからのデータインポート (4)プロジェクトの総合的推進 (1)リシーケンスによる臨床情報・ゲノム情報DBの構築 ①リシーケンスDBの構築 (1)標準SNP DBの構築 ①標準SNP DBの構築のための統計遺伝学手法の開発 (2)GWAS(ゲノムワイド関連解析)DBの構築 ①GWAS DBの統計遺伝学手法および解析ツールの開発 (1)GWAS DBの構築 ①GWAS 第一ステージ疾患関連SNP探索手法の研究開発 (2)リシーケンスによる臨床情報・ゲノム情報DBの構築 ①リシーケンスDBの解析手法の開発	(1)標準SNP、GWAS、CNV、リシーケンスのデータベースについて、合計40プロジェクト(studies)以上のデータを登録すること。 (2)これらのデータの受入れ・登録およびデータの提供が円滑に行われるよう、倫理社会的手続きおよび作業手順を整備すること。	(1) SNP control (2 studies), GWAS :18疾患 (25 studies)、17量的形質 (17 studies)、CNV control (1 study)、CNV case control :5疾患 (5studies)を内部DBに登録し、そのうち、GWAS:9疾患 (10studies)とSNP control (2 studies)、CNV control (1 study)は公開DBにも登録し、公開済み。残りのデータについては、論文のacceptとともに、順次公開していく。リシーケンスDBは、ALD、ALSについて一般の登録を新たに各1セット受け付けた。 (2) 倫理審査は、中核機関の協力も得て、データ共有委員会を発足し倫理社会的な手続き及び作業手順を整備し、http://gwas.lifesciencedb.jp/gwas/db_policy_en.htmlで公開するとともに、現在運用中。	(1) 新たに約5プロジェクトのデータを受け入れ・登録する (2) 学会・シンポジウム等で本DBをアピールし、認知度を上げる (3) 1(2)で定めた手順での、データの受け入れ・登録・データ提供の運用	(1) SNP control, GWAS 2008/8から、CNV control DB 2010/1から公開。CNV case control DBは、登録データの論文のacceptがまだなので現在非公開であるが、acceptされ次第公開。 (2) ALS resequence DBは2009/12、Parkinson resequence DBは2010/3、HSP、ALDについては、2010/5に公開済み。 (3) アクセス数(公開用のみの統計値): GWAS DB関連: 2009年 50,056 pages, 178,531 hits 2010年 53,234 pages, 224,809 hits Reseq DB関連: 2009年 4,637 pages 17,387 hits 2010年 13,113 pages 60,790 hits 内部DBについては、上記2.(1)のstudiesのデータを登録している研究グループの方々にご利用いただいている。 (4) データ数: GWAS DB関連: 上記2.(1)の通り Reseq DB関連: ・ ALS contents - 107 文献から638エントリーを抽出 (178 unique variants) familial ALS / sporadic ALS: 70% entries are related to familial ALS - 新たに実験的に発見した変異+臨床情報 - サブミッションしてもらった変異+臨床情報 ・ PD contents - 222 文献から1508エントリーを抽出 (472 unique variants) - 新たに実験的に発見した変異+臨床情報 ・ HSP contents - 103 文献から988エントリーを抽出(347 unique variants) ・ ALD contents - 49 文献から274エントリーを抽出(157 unique variants) - サブミッションしてもらった変異+臨床情報																																																																																

機関名		課題項目	目標と達成状況			
理研		①シロイヌナズナオミックス情報注釈 ②高等動物由来タンパク質構造のアノテーション ③微生物由来タンパク質構造のアノテーション ④アノテーションシステムの開発運用とデータ変換	①委託事業の9月末時点の判断基準になる目標 (サブテマ①) 文献情報から収集したシロイヌナズナ変異体の遺伝子および表現型の変異情報に基づき、表現型をオントロジーを用いて整理すると同時にそれぞれの表現型変異に対応する遺伝子(TAIRのAGIコード)を列挙する。9月末までに400種類の表現型を登録し、330の遺伝子コードへの対応付けを行い、これらすべてを理研サイネスを通じて提供する。 (サブテマ②) タンパク3000プロジェクトで解明された高等動物由来のタンパク質構造データのうち、累計3万1千件の回折実験データ(画像データ枚数)を公開し、それに関わる実験情報についても関連づけを完了させ、9月末までに可能な約50万件の結晶観察データについても提供する。 (サブテマ③) BacpediaのオリジナルDB(実験データを有するタンパク質約8700種類)を公開し、一般の研究者にとりて使いやすい環境を整える。 (サブテマ④) SciNeS上の検索システムの改良作業を進め、9月末までに植物関係の35のデータベースについて横断的な関連情報の検索機能を提供する。	②9月末時点の達成状況 (サブテマ①) シロイヌナズナミュータントの表現型情報を扱った論文266件を対象に、マニュアルキュレーションを実施し、表現型変異・その変異を観察した実験条件・関与が考えられる遺伝子などの情報を抽出した。これに基づき、438種類の表現型をオントロジーを用いた標準化し、370種類の遺伝子(AGIコード)との関連付けを行った。これらの結果をデータベース化し、サイネスを通じて公開した。 (サブテマ②) 上記1.の目標を達成した。 ・2009年5月: 回折実験データ2万件と付随する実験データを理研サイネス上に公開した。 ・2010年8月: 回折画像データ1万1千件と付随する実験データを理研サイネス上に公開した。 ・2010年9月: 約50万件の結晶観察データと付随する実験データを提供した。 (サブテマ③) BacpediaのオリジナルDB(実験データを有するタンパク質約8700種類)を公開し、一般の研究者にとりて使いやすい環境を整えた。 (サブテマ④) ・植物統合データベース( <a href="http://scinets.org/db/plant">http://scinets.org/db/plant</a> )を開発し、この下に植物関係のデータベースを統合化した。 ・SciNetS上でセマンティックリンクにより関連付けられた、遺伝子や表現型などのバイオデータアイテムを実用的速度で横断検索するWeb準拠のサービスを実現させた。	③上記達成状況を踏まえたプロジェクト終了までの目標 (サブテマ①) ・未処理の文献のキュレーションを進める。年度末までに270件ほどの文献を処理し、環境応答性や種子生産性に関する表現型情報を中心に標準化とデータベース化を進める。 ・シロイヌナズナ変異体データベースなどのDBと上記マニュアルキュレーションの成果との統合化をすすめる。植物オミックス研究者の意見を聞きながら、実験データ・文献データを問わずフェノーム情報を各種オミックス情報の検索が容易にできるようにシステムを整備する。 (サブテマ②) タンパク3000プロジェクトで解明された高等動物由来のタンパク質構造データ(X線による結晶構造解析)のうち、3万1千件の回折実験データ(画像データ枚数)を公開し、それに関わる実験情報についても関連づけを完了させ、可能なものについては公開する。 (サブテマ③) 理研播磨研究所における微生物由来のタンパク質構造データに付随する実験データに関し、統合データベースの一部として生命科学分野の幅広い研究者に役立つ体制を確立する。 (サブテマ④) シロイヌナズナで実施した統合化をモデルケースとして、表現型情報の標準化・上位オントロジーの整備を進め、シロイヌナズナ以外の生物種(マウス等)をも含めたデータベース統合化を実現する。	④成果の概要 (サブテマ①) ・植物表現型統合データベースをサイネス上に構築し、 <a href="http://scinets.org/item/cria224u4i">http://scinets.org/item/cria224u4i</a> から共有化した。 ・表現型の標準化は6種類の公開オントロジーへの意味リンクとして表現した。このDB構造は他のDBでも共通に使えるよう考慮して設計した。 (サブテマ②) 理研サイネスの下記URLに回折実験データ累計約3万1千件と付随する実験データを公開した。(結晶観察データ約50万件をテストサーバで編集。)URL: <a href="https://database.riken.jp/sw/links/ja/ria46i/">https://database.riken.jp/sw/links/ja/ria46i/</a> (サブテマ③) BacpediaオリジナルDBの公開画面(2010年7月6日公開、現在まで4280ページビュー) <a href="http://bacpedia.harima.riken.jp">http://bacpedia.harima.riken.jp</a> (サブテマ④) 各種オミックス間の関係を定義すべく新規に開発された「上位オントロジー」に準拠して、対象データベースを分類し直した。トップページに示すメニューにはこの分類体系を反映させている。 ・トップページから遺伝子を検索する場合、遺伝子データにキーワードが含まれない場合であっても遺伝子に関連付けられているキーワードを含む表現型データを見出し、当該遺伝子を推論し検索することができるようになった。また検索速度については、データアクセス種のチェックも含まれていた場合高々1秒で結果が表示されるまで高速化が実現できた。さらには、JSONPを介したキーワードを含むデータアイテム数を返すインターフェイスを構築し、ライブサイネス統合データベースの横断検索機能との連携が実現された。
	産総研糖鎖	①運営と開発体制の準備 ②データ提供機関との交渉 ③統合データベース構築	・腫瘍マーカーDB(TuMaRDB)の公開 ・グリコシダーゼの基質特異性の情報整備が公開間近になっていること ・糖鎖関連遺伝子の阻害剤DBが公開間近になっていること ・糖鎖関連特許DBが公開間近になっていること ・オンラインプロトコルDB(GlycoPOD)の公開、今年度分原稿待ち ・オンラインレポートの今年度分原稿待ち ・有機合成DBのデータ移行のシステム完了(検索システムを開発中になっていること) ・糖鎖関連遺伝子の発現情報解析システムの構築完了(データを入力して仮運用になっていること)	・腫瘍マーカーの精度をDB化し、公開できる状態。 ・グリコシダーゼの基質特異性の情報整備が公開間近である。 ・糖鎖関連遺伝子の阻害剤の構造を検索できるようにした。DBが公開間近である。 ・糖鎖関連特許DBが公開間近である。 ・オンラインプロトコルDB(GlycoPOD)のデータベースの説明文を待っている状況であり、公開が近い。今年度分原稿を執筆者に依頼済み。 ・オンラインレポートの今年度分原稿を依頼済み。3報の原稿が届いている。 ・有機合成DBのデータ移行のシステムが完成。構造検索エンジンも納品されて、データが渡ってくるのを待っている。 ・糖鎖関連遺伝子の発現情報解析システムがほぼ完成している。発現情報を解釈して部分的な構造を推測するツールの開発を行っている。	1. 平成21年度と平成22年度の2年間で100以上のプロトコルを収集する。年度末までに公開する。 2. グリコシダーゼのデータベースを構築する。 3. 糖鎖関連遺伝子のDNAマイクロアレイのデータを公開するシステムを構築 4. 糖鎖と病原菌のデータベース(PACDB)の更新 5. 糖鎖関連のマーカーの情報を収集しDB化する 6. 糖鎖関連の特許情報を集め整理しDBとして公開する 7. オンラインレポートの原稿を20日後にする 8. 直観的なポータルサイト実現のための統合検索システムの構築を実現する。できるだけDB間の連携を高められるようにする	横断・統合検索 キーワードによる横断検索 構造検索(CFGシンボル) 構造検索(JChemBaseを利用) 統合検索(プロトタイプ版を公開・開発中) 個別に開発したDB Lectinと糖鎖の結合情報のDB(LfDB) 糖鎖標準品の質量分析器で計測した際のスペクトルデータベース(GMDB) 糖タンパク質のデータベース(GlycoProtDB) 合成経路・化合物検索(開発中) グリコシダーゼ基質特異性DB(公開近い) GlycoGene阻害剤DB(GGIDB・公開近い) 糖鎖科学の実験プロトコルオンライン(GlycoPOD・公開近い) 糖鎖関連疾患とその原因遺伝子のデータベース(GDGB・未公開) Functional Glycomics with Knockout mice Database(未公開) 病原体と宿主の糖鎖との結合情報 DB(PACDB) JCGGB・オンラインレポート(一部公開・継続収集) 糖鎖関連特許のデータベース(未公開) 癌などに関する糖鎖マーカーのデータベース(TumaRDB・公開近い) 糖鎖関連遺伝子DNAマイクロアレイ・公開するためのデータベース(開発終了近い)
補完課題	①受付システムの高度化と運用 ②公開システムの高度化と運用 ③管理システムの高度化と運用 ④国際連携	□登録データからSRA ToolkitによりSRAファイルを作成しデータリストとともにDRAから公開する(9月末時点)	□SRA Toolkitにより登録データをSRAファイルに変換するシステムを整備した。公開日になると、SRAファイルから汎用されているfastqファイルを作成し、fastqとメタデータをDRAから公開している。同時にSRAファイルとメタデータはミラーリングのためNCBI Sequence Read Archiveに送付している。これにより、データの受付、作成、管理の一連の流れでNCBIに依存していた部分がなくなり、全ての処理をDRA自前でできるようになった。なお、NCBIに確認したところ「データリストは不要」とのことであったので、データリストは作成していない。 □メタデータ作成支援ツールMetaDefineに登録データのもととなるテンプレートを提供する機能を実装した(現在、Whole Genome Shotgun、TranscriptomeとBarcode)。 ・MetaDefineで、同一のランで由来サンプルをタグ配列で区別している「Barcodeデータ」のメタデータを作成可能にした。 □DDBJ Sequence Read Archiveの登録受付件数が336件に達した。 □DDBJ Sequence Read ArchiveとDDBJ Omics Archiveについての論文を公表した。 □Kodama Y et al(2010) Biological Databases at DNA Data Bank of Japan in the Era of Next-Generation Sequencing Technologies. Adv Exp Med Biol 680:125-135 □Sequence Read ArchiveへのDDBJ/EBI/NCBIの各種の取り組みをNucleic Acids Research誌に投稿した。 □NEDO 完全長cDNA構造解析プロジェクト「由来の波形データ」2,445,650件をDDBJ Trace Archiveへ登録した。本データはDDBJ Trace ArchiveとNCBI Trace Archiveから公開されている。	□メタデータの検索系、及び、それと連動したfastqファイルのダウンロード提供システムを完成させる(3月末時点)	DRAに登録された各種シーケンサ由来のランデータからSRA ToolkitによりSRAファイルを作成するシステムを整備した。 【従来】※青字がNCBI SRAに依存していた部分 登録データ → NCBI SRAにアップロード → SRAファイル作成 → SRAファイルとfastqファイル公開 → SRAからfastq取得、DRAで公開 【9月から】 登録データ → SRAファイル作成 → DRAでfastqファイルを作成、公開 → SRAファイルをNCBIにアップロード → NCBIから公開	
遺伝研	①受付システムの高度化と運用 ②公開システムの高度化と運用 ③管理システムの高度化と運用 ④国際連携	□登録データからSRA ToolkitによりSRAファイルを作成しデータリストとともにDRAから公開する(9月末時点)	□SRA Toolkitにより登録データをSRAファイルに変換するシステムを整備した。公開日になると、SRAファイルから汎用されているfastqファイルを作成し、fastqとメタデータをDRAから公開している。同時にSRAファイルとメタデータはミラーリングのためNCBI Sequence Read Archiveに送付している。これにより、データの受付、作成、管理の一連の流れでNCBIに依存していた部分がなくなり、全ての処理をDRA自前でできるようになった。なお、NCBIに確認したところ「データリストは不要」とのことであったので、データリストは作成していない。 □メタデータ作成支援ツールMetaDefineに登録データのもととなるテンプレートを提供する機能を実装した(現在、Whole Genome Shotgun、TranscriptomeとBarcode)。 ・MetaDefineで、同一のランで由来サンプルをタグ配列で区別している「Barcodeデータ」のメタデータを作成可能にした。 □DDBJ Sequence Read Archiveの登録受付件数が336件に達した。 □DDBJ Sequence Read ArchiveとDDBJ Omics Archiveについての論文を公表した。 □Kodama Y et al(2010) Biological Databases at DNA Data Bank of Japan in the Era of Next-Generation Sequencing Technologies. Adv Exp Med Biol 680:125-135 □Sequence Read ArchiveへのDDBJ/EBI/NCBIの各種の取り組みをNucleic Acids Research誌に投稿した。 □NEDO 完全長cDNA構造解析プロジェクト「由来の波形データ」2,445,650件をDDBJ Trace Archiveへ登録した。本データはDDBJ Trace ArchiveとNCBI Trace Archiveから公開されている。	□メタデータの検索系、及び、それと連動したfastqファイルのダウンロード提供システムを完成させる(3月末時点)	DRAに登録された各種シーケンサ由来のランデータからSRA ToolkitによりSRAファイルを作成するシステムを整備した。 【従来】※青字がNCBI SRAに依存していた部分 登録データ → NCBI SRAにアップロード → SRAファイル作成 → SRAファイルとfastqファイル公開 → SRAからfastq取得、DRAで公開 【9月から】 登録データ → SRAファイル作成 → DRAでfastqファイルを作成、公開 → SRAファイルをNCBIにアップロード → NCBIから公開	
九工大	①蛋白質と変異体の熱力学データベースの構築と統合 ②蛋白質・核酸相互作用の熱力学データベースの構築と統合 ③蛋白質・核酸相互作用データの生成と統合 ④XMLデータフォーマットやオントロジーなどの統合化技術の開発	・蛋白質熱力学データベースおよび蛋白質・核酸相互作用データベースのそれぞれ新規700件程度の熱力学データと構造データのクロスレファレンスを作成する。蛋白質・蛋白質相互作用データは格納用データベースシステムを完成させる。 ・蛋白質・核酸相互作用データをXML化し公開する。 ・統合DBセンターと連携して熱力学データの文献収集とデータ抽出の自動化技術の開発に必要な解析やツールの評価を行う。 9月末までに実用的な自動収集が行えるようにすることを目標とする。	・蛋白質熱力学および蛋白質・核酸相互作用のそれぞれ新規分データ約700件と、対応する構造データのクロスレファレンスを作成した。蛋白質・蛋白質相互作用データについては、格納用のデータベースのプロトタイプを作成し試験を行っている。 ・上記の蛋白質と核酸の相互作用の熱力学データの更新分についてXMLフォーマットに変換した。 ・熱力学データを含む文献を自動的に収集する目的でTogoDocをカスタマイズし、収集の効率が大幅に改善しほぼ実用の域に達した。	・蛋白質熱力学データおよび蛋白質・核酸相互作用データの収集を継続し、構造データとのクロスレファレンスを作成する。蛋白質・蛋白質相互作用データは格納用データベースシステムを完成させる。 ・これらのデータはXML化し公開する。 ・熱力学文献自動収集用にカスタマイズしたTogoDocは実用段階に達したので、今後実地の作業で使用し、こまかいチューニングを行いさらに使いやすいようにする。	蛋白質熱力学データベースおよび蛋白質・核酸相互作用データベースは、それぞれ25,000件、10,700件のデータを蓄積し、対応する構造データとのクロスレファレンスを作成し統合化した。類似のデータベースが存在しないため、この分野では重要なリソースとして世界中の研究者に利用されている。1月～7月のアクセス(ページリクエスト数)は約46,000回となっている。これらのデータベースはすでに200件以上の論文に引用され、データベースを用いて解析を行った研究論文も数多く出版されている。 データベース構築で最も時間のかかるのが熱力学データを含む文献の収集とデータの抽出であるが、TogoDocをカスタマイズすることにより、文献収集の効率を大幅に改善することができた(DBCLSとの共同研究)。	

表2 アクセス件数と転送データサイズ

	2010年1月	2月	3月	4月	5月	6月	7月	8月	9月
アクセス件数	1684	1459	1886	2703	2642	2475	2196	2972	2758
バイト数(MB)	146	133	193	784	345	335	315	364	380

	2010年1月	2月	3月	4月	5月	6月	7月	8月	9月
アクセス件数	468	692	333	594	440	1530	6399	10969	564
バイト数(MB)	3177	147	2719	6818	3898	31472	575548	33922	2931

7月と8月に機械的なデータダウンロードアクセスがあり、件数が急増した