

独立行政法人科学技術振興機構
バイオインフォマティクス推進センター事業

生命情報データベースの高度化・標準化

第Ⅱ期 研究開発課題

(追跡調査報告書)

研究開発課題：マルチモーダル統合バイオDB

(平成18年4月～平成23年3月)

代表研究者氏名：森下 真一

(東京大学大学院新領域創成科学研究科 教授)

目次

本編

1. 研究開発による成果.....	1
1.1. 研究開発課題の目標及び新規性・独創性	1
1.2. 研究開発開始時の課題を取り巻く研究・技術水準及び分野における 課題の位置づけ	1
1.3. 研究開発終了時の成果概要	2
1.4. 研究開発の達成度	3
2. 研究開発による成果の活用状況や発展状況.....	3
2.1. 構築されたデータベース・ソフトウェア等の活用状況.....	3
2.1.1. データベース・ソフトウェア等の継続状況や発展状況.....	3
2.1.2. 第三者によるデータベース・ソフトウェア等の活用事例	4
2.1.3. データベース・ソフトウェア等へのアクセス数／ダウンロード数、visit 数	5
2.2. 課題終了後の研究開発成果の継続状況や発展状況.....	9
2.2.1. BIRD 終了から現在まで、BIRD で取り組んだ課題に関連した研究開発成果の 継続状況や発展状況（国内外の研究者との共同研究へ発展した等）	9
2.2.2. BIRD 終了後に発表された論文リスト.....	11
2.2.3. 研究開発成果の現在の国際的な評価・位置づけ	11
2.3. 現在の科学技術研究における研究開発成果の活用状況、発展状況のまとめ.....	12
3. 研究開発による成果の科学技術的、社会的及び経済的な効果.....	13
3.1. バイオインフォマティクス分野及びライフサイエンス分野の研究への貢献.....	13
3.1.1. バイオインフォマティクス分野の研究への貢献	13
3.1.2. ライフサイエンス分野（バイオインフォマティクス以外）や その他、科学技術分野の研究への貢献	13
3.2. 人材育成の面から参加研究者の活動状況	14
3.2.1. 研究開発に参加した研究者のキャリアアップ	14
3.3. 社会的及び経済的な効果	14
3.3.1. 研究開発成果が大学や公的研究機関、企業等で、応用に向けて 継承または発展した例	14
3.3.2. その他、研究開発成果が社会的、経済的な効果・効用につながる兆し、 可能性	14

資料編

1. 論文リスト	16
2. 主要論文の被引用回数.....	20

3.	学会招待講演・基調講演.....	23
4.	新聞発表等	24
5.	特許出願・成立	24
6.	学会賞等の受賞	24
7.	グラントの獲得実績.....	24
8.	書籍等の執筆実績	25
9.	総説の執筆実績	25
10.	参加研究者の活動状況.....	25

本調査報告書は平成 26 年 4 月に作成
表紙の代表研究者所属は平成 25 年 10 月時点

1. 研究開発による成果

1.1. 研究開発課題の目標及び新規性・独創性

DNA 配列、遺伝子発現量、バイオイメージ、パスウェイ、文献などの様々なバイオデータが急速に蓄積されてきている。膨大なデータから目的の情報を抽出するために BLAST、クラスタリング、PubMed などのソフトウェアが利用されるが、これらのソフトウェアは単一のデータタイプを処理することに特化している。現実には単独のデータタイプだけを利用するよりも、むしろ異なるデータタイプのバイオデータを組み合わせ、総合的に活用することで遺伝子情報を絞り込み、効率の良い実験計画立案のために利用することが多い。

本研究では、ゲノム、トランスクリプトーム、プロテオーム、フェノームといった異なるタイプのバイオデータを統合し、効率よく活用するための統合化システムであるマルチモーダル統合バイオ DB を実現することを目標とした。多様な生物種のマルチモーダル統合バイオ DB 構築を可能にするツールキットの開発により、出芽酵母に加えて、メダカ、ショウジョウバエ、カイコ、ヒトへと対象を広げ、マルチモーダル統合バイオ DB サーバーの構築を目指した点に新規性が認められる。また、超並列クラスター型計算機を利用したワークフローシステムを開発することで、超高速シーケンサーからの大量配列データの処理を可能とする実用的マルチモーダル統合バイオ DB の実現を目指した点も新規・独自性がある。

1.2. 研究開発開始時の課題を取り巻く研究・技術水準及び分野における課題の位置づけ

遺伝子情報や蛋白質の配列・構造に関する情報などが日々蓄積され、巨大なデータベースが構築されている。生物学に利用されるデータベースとしては、①DNA、RNA の塩基配列とその説明、文献情報などを記載した「塩基配列データベース」(GenBank、EMBL、DDBJ)、②アミノ酸の配列情報とその説明、文献情報などを記載した「アミノ酸配列データベース」(PIR、Swiss-Prot、PRF)、③タンパク質相互作用に関与する部位の特徴的な配列や立体構造パターン (モチーフ) や、タンパク質のファミリー分類を登録した「タンパク質機能データベース」(PROSITE、Pfam、Motif)、④タンパク質立体構造データと構造を分類・比較したデータや、フォールドを登録した「タンパク質立体構造データベース」(PDB、SCOP、CATH)、⑤複数分子間の相互作用ネットワークを集積した「パスウェイデータベース」(KEGG)、⑥組織、細胞における遺伝子発現プロファイルを登録した「遺伝子発現データベース」(LSBM、ArrayExpress)、⑦代謝マップや代謝産物を登録した「メタボロームデータベース」、⑧学術論文を集めた「文献データベース」(PubMed、Google Scholar) などがある。大量のデータから科学的分析を迅速に実施するには、マルチモーダルデータの表現形式を定式化し、データ変換操作を代数的に組み合わせることができるよう理論基盤を作ることが解決につながると考えられたが、多様なデータ構造をもつマル

チモーダルデータを扱った先行研究例は少なかった。

異なるデータタイプのバイオデータを総合的に活用するシステムとして、出芽酵母では既に国際的研究グループ YSBN が存在したが、他の様々な生物種においてもマルチモーダル統合バイオ DB の構築が強く求められていた。本研究開発課題においては、出芽酵母に加えて、発生遺伝学のモデルとして利用されるメダカ、ショウジョウバエ、害虫研究のモデルとしてカイコ、そしてヒトへと対象生物種を広げ、異なるデータタイプの大量のデータを処理できるデータベースシステムの構築を行った。データベースシステムの構築は、計算資源とソフトウェアの最適化により対応可能であったが、データベースの構築と公開だけで終わることなく、マルチモーダル統合バイオ DB の考え方が生物現象の理解・解明に役立つことを示すため、マルチモーダルデータを分析・活用することにより新たな科学的知見を提示することも目指した。

1.3. 研究開発終了時の成果概要

- 1) UTGB Toolkit の研究開発により、多様な生物種のマルチモーダル統合バイオ DB サーバーの構築が可能になった。ゲノム、インターラクトーム、トランスクリプトーム、フェノームデータのマルチモーダル統合 DB を、ヒト (PrimerStation、siDirect、dsCheck)、出芽酵母 (UTGB/yeast、SCMD)、ショウジョウバエ (MachBase、UTGB/fruitfly)、メダカ (UTGB/medaka)、カイコ (UTGB/silkworm) を対象に構築・公開した。特定領域研究ゲノム等のアカデミア研究者らにより収集された新しいデータも随時追加・更新・メンテナンスした。

また、ゲノムブラウザー UTGB の公開により、初心者でも個人用にゲノムブラウザーを構築でき、データベースの移植性向上 (SQLite JDBC ライブラリによる研究開発)、シェル機能の開発 (UTGB Shell) によるデータベース管理の効率化が可能になった。

- 2) 当初の研究計画では予想されていなかった超高速シーケンシング技術に対応するため、マルチモーダル統合バイオ DB の研究計画を前向きに軌道修正した。超高速シーケンサーの塩基配列解読量の増加に対応するため、多数の CPU を使用してデータ処理を並列化した並列計算機を使用し、マルチモーダル統合バイオ DB の並列化を試みた。これにより、以下のマルチモーダルデータの高速な収集が可能になった。

- ・ エピゲノム情報 (DNA メチル化、ヌクレオソーム位置の分布、ヒストン修飾)
- ・ SAGE 法を拡張した定量的転写開始時点情報、広いダイナミックレンジの遺伝子発現情報
- ・ 全長 cDNA 配列決定
- ・ 遺伝子多型

- 3) DNA の三次元構造（ヌクレオソーム構造）が DNA の変異に相関することを明らかにした。具体的には、超高速シーケンサーを活用し、メダカのヌクレオソーム構造を DNA 全体にわたって網羅的に分析し、その結果、遺伝子の転写開始点の下流におけるヌクレオソーム構造が DNA 変異の周期性を引き起こす要因となっていることを示した。本研究により、遺伝的多様性が生じる過程の一端を明らかにした。

1.4. 研究開発の達成度

開始当初のデータ生産速度であれば、当初計画で十分対応できたが、超高速シーケンサーの普及に伴い、ゲノム、トランスクリプトーム（RNA・Seq、転写開始点）の収集速度は指数関数的に増加した。例えば、1日あたりの塩基配列解読量は、2002年200万塩基、2007年2億塩基である。さらに、ゲノム、mRNAだけでなく、従来は情報収集が困難であったエピゲノムデータ（DNAメチル化、ヒストン修飾、ヌクレオソーム構造、大域的クロマチン構造など）や遺伝子発現量解析データも収集可能になり、超高速シーケンサーは、単なるゲノム解読装置ではなく、「多様な生物現象を観測する装置」となった。これら新しい情報は全てマルチモーダルなデータであり、研究対象として積極的に対応しなければならないものであった。こうした解読量の飛躍的増加や、新しい情報の登場に対応するため毎年研究計画を柔軟に変更し、前倒しに対応し迅速な研究開発を行うことにより、当初の研究開発計画を超えライフサイエンス分野の情報基盤整備に大きく貢献した。また、マルチモーダルな統合検索により新たな生物学的発見が可能であることを実証し、ライフサイエンス分野における新たな研究スタイルを確立し、レベルの高い成果を出したことは特筆に値する。

2. 研究開発による成果の活用状況や発展状況

2.1. 構築されたデータベース・ソフトウェア等の活用状況

2.1.1. データベース・ソフトウェア等の継続状況や発展状況

	ツール名	URL	公開状況	備考
(1)	UTGB Toolkit (University of Tokyo Genome Browser)	http://utgenome.org/	公開中（無料）	
(2)	SCMD	http://scmd.gi.k.u-tokyo.ac.jp/datamine/	公開中（無料）	

(3)	siDirect/dsCheck	http://sidirect2.rnai.jp/	公開中（無料）	
(4)	UT Genome Browser (Medaka)	http://utgenome.org/medaka/	公開中（無料）	
(5)	UT Genome Browser (Yeast)	http://yeast.utgenome.org/	公開中（無料）	
(6)	PrimerStation	http://ps.cb.k.u-tokyo.ac.jp/index.html	公開中（無料）	
(7)	MachiBase (UT Genome Browser / Fruitfly)	http://machibase.gi.k.u-tokyo.ac.jp/	公開中（無料）	
(8)	FreClu (Frequency-based de novo short read Clustering)	http://mlab.cb.k.u-tokyo.ac.jp/~quwei/DeNovoShortReadClustering/	公開中（無料）	
(9)	sqlite-jdbc	http://mvnrepository.com/artifact/org.xerial/sqlite-jdbc	公開中（無料）	SQLデータベースを動かすライブラリ。ゲノム分野以外にも広く使われている。
(10)	snappy-java	http://mvnrepository.com/artifact/org.xerial.snappy/snappy-java	公開中（無料）	データ圧縮ライブラリ。ゲノム分野以外にも広く使われている。

2.1.2. 第三者によるデータベース・ソフトウェア等の活用事例

- (1) UTGB Toolkit (University of Tokyo Genome Browser) : maven.utgenome.org から利用状況に応じてソフトウェアライブラリがダウンロードされる仕組みになっている。
- (2) SCMD : SGD（世界最大級の酵母ゲノムのデータベースサイト : <http://www.yeastgenome.org/>）より、SCMD で公開している遺伝子破壊株画像・解析結果へのリンクがある。たとえば本アプローチはハーバード大学の Altschuler らにより薬剤応答の指標として活用されている (Perlman et al. *Science*, 2004)。また、このサーバーデータから得られる情報を解析した我々の論文 (Ohya et al. *PNAS*, 2005) は 2012 年以降 41 回引用されている (Google scholar より)。
- (3) siDirect/dsCheck : siDirect は 2012 年以降に siRNA 関連の学術論文 35 件に引用されている。dsCheck は 2012 年中に RNAi 関連の学術論文 6 件に引用されている。
siDirect に関する論文では、東京医科歯科大学、慶応大学、RNAi 社等の研究者の共著で子宮頸がん治療用の siRNA 設計に利用したことが記載されている。また、RNAi 社

HP には、東京大学西郷薫教授、程久美子准教授、森下真一教授等のウェット、ドライ研究成果を基盤に、創薬研究に必須の機能を強化した RNAi 配列設計システムとして Enhanced siDirect®を利用していることが記載されている。

また Xi Chen らの論文(PLOS ONE 2013)において、siRNA 設計における”popular tool”として siDirect が記載されている。

- (6) PrimerStation : 2012 年 5 月に出版された Real-time PCR を解説した書籍に引用されている。Quantitative Real-time PCR in Applied Microbiology, Caister Academic Press (May 8, 2012)
- (7) MachiBase (UT Genome Browser/Fruitfly) : Ting Ni et al. (Nature Methods vol 7. 2010)に MachiBase のデータを利用した論文が報告されている。真核生物の複雑な転写開始をマッピングするペアエンドシーケンシングにおいて、ペアエンド TSS 分析法が強力なツールになることを示した論文で、MachiBase による結果が比較例として利用されている。同様の例として、(Nechaew et al., *Science* 2009) がある。
- (8) FreClu (Frequency-based de novo short read Clustering) : 2012 年中にシーケンシングデータ処理関連の学術論文 7 件に引用されている。Suffix array の線形時間構築で知られる Aluru 教授がまとめた次世代シーケンサーのためのエラー修正アルゴリズムのサーベイの中で、パイオニア的アルゴリズムとして紹介されている (Aluru et al. Briefings in bioinformatics 2013)。
- (9) sqlite-jdbc, snappy-java のライブラリは Maven Central (Java プログラムのための世界共通公開レポジトリ) にて配布されており、どちらも月 2 万を超えるダウンロードがある。

2.1.3. データベース・ソフトウェア等へのアクセス数／ダウンロード数、visit 数

	ツール名	アクセス数、ダウンロード数、visit 数	集計期間	備考
(1)	UTGB Toolkit	アクセス数 : 4,303,762 visitor 数 : 219,981 ((4) UT Genome Browser Medaka と合算した数値)	2012 年 11 月 ～2013 年 10 月 (1 年間)	月ごとの値は 表 1a を参照
	UTGB Toolkit (maven.utgenome.org での公開分)	アクセス数 : 564,980 visitor 数 : 82,219	2012 年 11 月 ～2013 年 10 月 (1 年間)	月ごとの値は 表 1b を参照
(2)	SCMD	アクセス数 : 433,026 visitor 数 : 40,340	2012 年 11 月 ～2013 年 10 月 (1 年間)	月ごとの値は 表 3 を参照

(3)	siDirect/dsCheck	siDirect アクセス数：84,823 visitor 数：5,317 dsCheck アクセス数：3,979 visitor 数：1,463	2012 年 8 月 31 日 ～2013 年 3 月 (約 6 か月間)	月ごとの値は 表 4 を参照
(4)	UT Genome Browser (Medaka)	アクセス数：4,303,762 visitor 数：219,981 (1) UTGB Toolkit と合算 した数値)	2012 年 6 月 ～2013 年 5 月 (1 年間)	
(5)	UT Genome Browser (Yeast)	アクセス数：27,832 visitor 数：189	2012 年 6 月 ～2013 年 5 月 (1 年間)	月ごとの値は 表 5 を参照
(6)	PrimerStation	アクセス数：34,246 visitor 数：12,376	2012 年 4 月 ～2013 年 3 月 (1 年間)	月ごとの値は 表 6 を参照
(7)	MachiBase	アクセス数：20,271 visitor 数 3,073	2012 年 6 月 ～2013 年 5 月 (1 年間)	月ごとの値は 表 7 を参照
(8)	FreClu	アクセス数：3,401 visitor 数：1,721	2012 年 4 月 ～2013 年 5 月 (1 年間)	月ごとの値は 表 8 を参照
(9)	sqlite-jdbc	ビジター数：91,160 ページビュー202,333	2012 年 10 月 ～2013 年 9 月 (1 年間)	月ごとの値は 図 1 を参照
(10)	snappy-java	ビジター数：21,447 ページビュー：44,948	2012 年 10 月 ～2013 年 9 月 (1 年間)	月ごとの値は 図 1 を参照

表 1a. utgenome.org

Summary by Month										
Month	Daily Avg				Monthly Totals					
	Hits	Files	Pages	Visits	Sites	KBytes	Visits	Pages	Files	Hits
Oct 2013	11797	9115	3302	857	1943	1439826	15768	79261	218782	283132
Sep 2013	14110	10397	2873	846	2477	2463247	25383	86210	311936	423301
Aug 2013	16046	11153	3682	846	2364	2544869	26249	114148	345770	466441
Jul 2013	11416	8501	3158	631	2473	2096197	19591	97921	263542	353921
Jun 2013	11880	8905	2969	505	2357	1906874	15170	89071	267159	356419
May 2013	12796	9266	2961	531	2082	1810162	16466	91796	287259	396705
Apr 2013	12327	8958	3099	657	2248	1719255	19715	92992	268742	369832
Mar 2013	11753	8634	2580	483	2382	1686907	14983	79379	267681	364348
Feb 2013	10827	8132	2455	487	1884	1495648	13645	68754	227701	303167
Jan 2013	7956	6123	2076	453	1731	1249128	14061	64359	189831	246645
Dec 2012	12731	10047	3562	605	1645	2493880	18773	110429	311474	394663
Nov 2012	11506	8738	2871	672	1690	1697122	20177	86130	262141	345188
Totals						22603115	219981	1060450	3222018	4303762

表 1b. maven.utgenome.org

Summary by Month										
Month	Daily Avg				Monthly Totals					
	Hits	Files	Pages	Visits	Sites	KBytes	Visits	Pages	Files	Hits
Oct 2013	1576	889	635	226	451	2767404	5435	15250	21346	37827
Sep 2013	2035	1053	689	252	486	3813134	7585	20696	31613	61055
Aug 2013	1768	702	428	142	469	3100475	4420	13271	21790	54814
Jul 2013	1530	555	358	165	432	2138446	5126	11112	17216	47451
Jun 2013	1693	670	434	162	417	1834289	4870	13033	20103	50792
May 2013	1633	698	460	156	442	2587886	4846	14286	21638	50633
Apr 2013	2145	1222	771	227	648	3373885	6833	23131	36670	64351
Mar 2013	2559	804	514	185	813	4975554	5746	15950	24932	79356
Feb 2013	1359	790	551	237	662	969854	6659	15443	22145	38058
Jan 2013	826	693	478	272	561	2041170	8432	14847	21492	25632
Dec 2012	873	686	515	349	427	2668893	10832	15987	21275	27084
Nov 2012	930	702	576	381	393	1628576	11435	17288	21073	27927
Totals						31899566	82219	190294	281293	564980

表 2. scmd.gi.k.u-tokyo.ac.jp

Summary by Month										
Month	Daily Avg				Monthly Totals					
	Hits	Files	Pages	Visits	Sites	KBytes	Visits	Pages	Files	Hits
Oct 2013	557	438	146	96	878	239126	2318	3512	10523	13375
Sep 2013	1616	1376	260	121	1291	1002119	3648	7800	41302	48481
Aug 2013	2166	1959	443	111	1159	983169	3455	13757	60749	67162
Jul 2013	878	726	228	110	1056	582930	3414	7078	22529	27225
Jun 2013	935	796	170	111	1108	512747	3359	5128	23898	28056
May 2013	587	448	155	105	1189	695233	3281	4822	13901	18201
Apr 2013	865	650	200	114	1396	448824	3448	6001	19511	25978
Mar 2013	739	552	204	116	1337	475520	3625	6349	17140	22939
Feb 2013	1375	1170	268	120	1117	471064	3360	7525	32775	38503
Jan 2013	1553	1258	414	120	1353	724004	3745	12840	39027	48167
Dec 2012	1429	1141	391	110	1099	930155	3410	12130	35390	44327
Nov 2012	1687	1403	336	109	1052	759640	3277	10096	42107	50612
Totals						7824531	40340	97038	358852	433026

表 3. siDirect(<http://sidirect2.rnai.jp>)

Summary by Month										
Month	Daily Avg				Monthly Totals					
	Hits	Files	Pages	Visits	Sites	KBytes	Visits	Pages	Files	Hits
Mar 2013	453	396	118	29	617	442405	928	3660	11968	14058
Feb 2013	307	269	77	20	396	255341	571	2158	7532	8605
Jan 2013	402	360	95	24	513	446956	746	2962	11190	12475
Dec 2012	405	356	120	23	512	166843	735	3739	11036	12581
Nov 2012	408	380	111	24	478	199694	743	3339	11401	12256
Oct 2012	397	353	102	27	475	229278	841	3192	10959	12309
Sep 2012	407	354	100	24	487	413924	740	3019	10625	12231
Aug 2012	308	265	46	13	17	3658	13	46	265	308
Totals						2148099	5317	22115	74996	84823

表 4. dsCheck (<http://dscheck.rnai.jp>)

Summary by Month										
Month	Daily Avg				Monthly Totals					
	Hits	Files	Pages	Visits	Sites	KBytes	Visits	Pages	Files	Hits
Mar 2013	15	9	9	9	162	81	292	301	292	483
Feb 2013	17	8	8	8	161	83	241	251	248	483
Jan 2013	34	9	10	8	163	198	266	314	255	1071
Dec 2012	16	6	7	6	115	93	195	218	192	526
Nov 2012	19	5	6	5	116	108	169	191	177	599
Oct 2012	16	5	6	5	120	94	168	187	171	526
Sep 2012	9	4	4	4	99	48	125	132	127	283
Aug 2012	8	7	7	7	7	1	7	7	7	8
Totals						705	1463	1601	1499	3979

表 5. UT Genome Browser (Yeast)
(<http://yeast.utgenome.org>)

Summary by Month										
Month	Daily Avg				Monthly Totals					
	Hits	Files	Pages	Visits	Sites	KBytes	Visits	Pages	Files	Hits
Mar 2013	255	255	78	1	10	14507	38	2434	7922	7922
Feb 2013	173	173	38	1	9	2699	47	1070	4846	4846
Jan 2013	173	173	42	1	9	4384	38	1305	5369	5369
Dec 2012	128	128	20	0	8	2171	18	630	3990	3990
Nov 2012	116	116	14	0	4	3880	13	247	1975	1975
Oct 2012	155	155	40	1	6	973	17	573	2173	2173
Sep 2012	119	119	7	1	2	676	3	23	358	358
Aug 2012	0	0	0	0	0	0	0	0	0	0
Jul 2012	0	0	0	0	0	0	0	0	0	0
Jun 2012	42	42	11	0	4	455	15	335	1199	1199
Totals						29745	189	6617	27832	27832

表 6. PrimerStation
(<http://ps.cb.k.u-tokyo.ac.jp/index.html>)

Summary by Month										
Month	Daily Avg				Monthly Totals					
	Hits	Files	Pages	Visits	Sites	KBytes	Visits	Pages	Files	Hits
Mar 2013	110	47	78	45	732	13536	1404	2441	1473	3415
Feb 2013	91	36	65	42	591	10247	1195	1838	1017	2551
Jan 2013	101	39	70	42	692	15557	1318	2182	1223	3154
Dec 2012	86	38	61	41	654	12040	1288	1906	1180	2667
Nov 2012	82	38	50	32	586	13213	967	1519	1147	2474
Oct 2012	88	42	58	33	598	14320	1023	1803	1330	2749
Sep 2012	155	105	119	29	556	14779	896	3573	3157	4663
Aug 2012	64	27	40	25	462	7429	785	1264	843	1964
Jul 2012	58	24	39	26	463	7769	821	1227	774	1810
Jun 2012	64	28	38	26	524	8436	807	1150	860	1945
May 2012	143	34	112	29	545	11618	928	3499	1076	4444
Apr 2012	79	34	49	30	547	11546	924	1485	1042	2390
Totals						140491	12376	23887	15122	34246

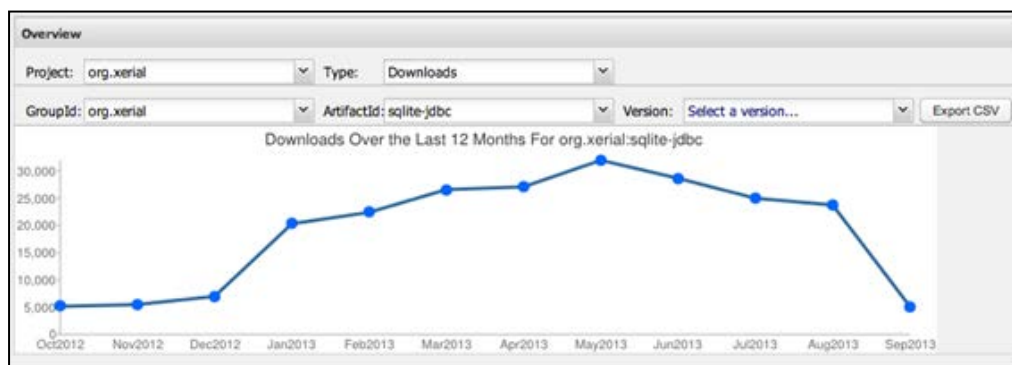
表 7. machbase.gi.k.u-tokyo.ac.jp

Summary by Month										
Month	Daily Avg				Monthly Totals					
	Hits	Files	Pages	Visits	Sites	KBytes	Visits	Pages	Files	Hits
Oct 2013	34	33	21	4	69	6160	101	518	801	827
Sep 2013	48	45	29	6	142	10987	184	886	1372	1458
Aug 2013	47	45	27	6	146	12407	215	849	1414	1465
Jul 2013	58	57	38	8	143	12627	273	1184	1783	1801
Jun 2013	70	66	40	9	160	18687	296	1222	1990	2118
May 2013	70	68	50	13	188	14699	422	1580	2138	2177
Apr 2013	62	57	43	11	149	9927	345	1290	1727	1873
Mar 2013	73	72	47	10	145	16267	316	1466	2253	2280
Feb 2013	49	48	27	6	105	11400	178	764	1366	1397
Jan 2013	45	42	23	5	108	14201	165	731	1324	1423
Dec 2012	67	60	43	12	161	13393	374	1344	1890	2079
Nov 2012	60	59	33	6	114	15838	204	992	1798	1823
Totals						156593	3073	12826	19856	20721

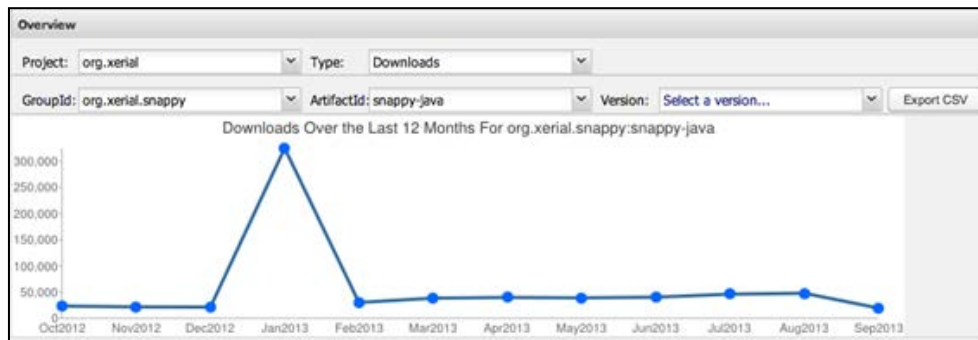
表 8. FreClu
(<http://mlab.cb.k.u-tokyo.ac.jp/~quwei/DeNovoShortReadClustering/>)

Summary by Month										
Month	Daily Avg				Monthly Totals					
	Hits	Files	Pages	Visits	Sites	KBytes	Visits	Pages	Files	Hits
Mar 2013	10	7	7	6	124	780181	188	246	242	339
Feb 2013	8	6	6	5	100	2783880	145	191	171	238
Jan 2013	7	5	5	4	102	3342139	139	183	172	245
Dec 2012	7	5	5	4	89	1980781	126	171	168	231
Nov 2012	18	7	6	4	130	1337460	146	193	233	544
Oct 2012	5	4	4	3	97	3677235	112	125	141	180
Sep 2012	7	4	5	3	109	4163962	119	153	145	211
Aug 2012	6	4	3	3	88	1771689	95	118	154	188
Jul 2012	7	6	4	3	96	4559033	117	134	192	235
Jun 2012	15	13	10	8	130	6976689	265	300	414	453
May 2012	9	8	6	5	104	9509278	175	191	250	285
Apr 2012	9	7	3	3	84	16751803	94	107	216	252
Totals						57634131	1721	2112	2498	3401

図 1. sqlite-jdbc 及び snappy-java の Maven Central での月別ダウンロード数



sqlite-jdbc



snappy-java

2. 2. 課題終了後の研究開発成果の継続状況や発展状況

2. 2. 1. BIRD 終了から現在まで、BIRD で取り組んだ課題に関連した研究開発成果の継続状況や発展状況（国内外の研究者との共同研究へ発展した等）

1) 開発した DB の継続状況

BIRD プロジェクト終了後サーバーの維持・管理のため、メンテナンスのしっかりした海外の著名なポータルサイト（UCSC、Ensembl、WormBase、SCD、NCBI SRA）にもデータを提供し、データが消滅することを防ぎ、万が一の際にもユーザーに支障が出ないようにしている。

2) 東京大学大学院理学系研究科 武田洋幸研究室との共同研究

メダカのゲノムデータベースに関して共同研究を継続しており、2 系統のメダカについて肝臓・胞胚・精巣のゲノムメチル化パターンを解読し公表している（Medaka Methylome Browser、<http://utgenome.org/medaka/methylome/>）。なお武田研究室とは JST の支援を受け、発生過程における三次元クロマチン構造の動態を観測分析するプロジェクトも継続している（JST CREST「DNA3 次元クロマチン動態の理解と予測」H25-30 年度）。

発表論文

Wei Qu, Shin-ichi Hashimoto, Atsuko Shimada, Yoichiro Nakatani, Kazuki Ichikawa, Taro L Saito, Katsumi Ogoshi, Kouji Matsushima, Yutaka Suzuki, Sumio Sugano, Hiroyuki Takeda, Shinichi Morishita. Genome-wide genetic variations are highly correlated with proximal DNA methylation patterns. Genome Research 22:1419-1425 (2012)

3) スタンフォード大学 Andrew Fire 研究室との共同研究

線虫の転写開始点について共同研究を継続し、その成果が線虫の転写開始点に関する情報

サイト (<http://wormtss.utgenome.org/>) として公表されている。なお Fire 研究室とは線虫ゲノムの再解読等でも共同研究を継続中である。

発表論文

Taro L. Saito, Shin-ichi Hashimoto, Sam Guoping Gu, J. Jason Morton, Michael Stadler, Thomas Blumenthal, Andrew Fire, and Shinichi Morishita. The Transcription Start Site Landscape of *C. elegans*. *Genome Research* 23(8), 2013, 1348-1361

4) 東京大学大学院医学系研究科 辻省次研究室との共同研究

「文部科学省 新学術領域研究 生命科学系 3 分野支援活動『ゲノム支援』」の枠組みの中で、国内の約 30 の研究室をゲノムタイピング・リシーケンシング及びゲノム医学インフォマティクスに関する支援を実施している。具体的には、次世代シーケンサー Illumina HiSeq2500 を使い、疾患関連遺伝子を、家系データを使って絞り込んでいる。これに関して下に掲げた論文が出ているほか、今後も公表を予定している。

発表論文

Koichiro Doi, (9 名), Shoji Tsuji, Shinichi Morishita. Rapid detection of expanded short tandem repeats in personal genomics using hybrid sequencing. *Bioinformatics* (2013) doi: 10.1093/bioinformatics/btt647

Kazuki Ichikawa and Shinichi Morishita. A simple but powerful heuristic method for accelerating k-means clustering of large-scale data in life science. *IEEE/ACM TCBB* (招待論文)

Yuji Takahashi, (37 名) Shinichi Morishita, Shoji Tsuji. Mutations in ERBB4 which disrupt the neuregulin-ErbB4 pathway cause amyotrophic lateral sclerosis type 19. *American Journal of Human Genetics, Am J Hum Genet.* 93(5), 900-5, 2013

Isojima T, Doi K, Mitsui J, Oda Y, Tokuhiko E, Yasoda A, Yorifuji T, Horikawa R, Yoshimura J, Ishiura H, Morishita S, Tsuji S, Kitanaka S. A recurrent de novo FAM111A mutation causes Kenny-Caffey syndrome type 2. *J Bone Miner Res.* 2013

Ichikawa Y, Ishiura H, Mitsui J, Takahashi Y, Kobayashi S, Takuma H, Kanazawa I, Doi K, Yoshimura J, Morishita S, Goto J, Tsuji S. Exome analysis reveals a Japanese family with spinocerebellar ataxia, autosomal recessive 1. *J Neurol Sci.* 2013 Aug 15;331(1-2):158-60

Shin-ichi Hashimoto, Katsumi Ogoshi, Atsushi Sasaki, Jun Abe, Wei Qu, Yoichiro Nakatani, Budrul Ahsan, Kenshiro Oshima, Francis HW Shand, Akio Ametani, Yutaka Suzuki, Shuichi Kaneko, Takashi Wada, Masahira Hattori, Sumio Sugano, Shinichi Morishita, Kouji Matsushima. Coordinated changes in DNA methylation in antigen-specific memory CD4 T cells. *Journal of Immunology*, 190(8):4076-91 (2013)

Ishikawa T, Okada T, Ishikawa-Fujiwara T, Todo T, Kamei Y, Shigenobu S, Tanaka M, Saito TL, Yoshimura J, Morishita S, Toyoda A, Sakaki Y, Taniguchi Y, Takeda S, Mori K. ATF6 α / β -mediated adjustment of ER chaperone levels is essential for development of the notochord in medaka fish. *Mol Biol Cell*. 24(9):1387-1395. (2013)

Atsushi Ishii, Yoshiaki Saito, Jun Mitsui, Hiroyuki Ishiura, Jun Yoshimura, Hidee Arai, Sumimasa Yamashita, Sadami Kimura, Hirokazu Oguni, Shinichi Morishita, Shoji Tsuji. "Identification of ATP1A3 Mutations by Exome Sequencing as the Cause of Alternating Hemiplegia of Childhood in Japanese Patients." *PLOS ONE* 8.2 (2013)

5) Pacific Biosciences 社 (PB 社) との共同研究

PB 社創業者の Steve Turner、Jonas Korlach と、次世代シーケンサー PacBio RS II を使って 10 Kb 以上の領域の DNA メチル化を推測する共同研究を行い論文投稿中である。この研究によって、従来は DNA メチル化の判定が困難だったリピート領域およびハプロタイプの DNA メチル化推定が初めて可能になった。

2.2.2. BIRD 終了後に発表された論文リスト

(2.2.1 で挙げたものを除く)

Funakoshi M, Tsuda M, Muramatsu K, Hatsuda H, Morishita S, Aigaki T. A gain-of-function screen identifies wdb and lkb1 as lifespan-extending genes in *Drosophila*. *Biochem Biophys Res Commun*.405(4):667-72 (2011)

Ogoshi K, Hashimoto S, Nakatani Y, Qu W, Oshima K, Tokunaga K, Sugano S, Hattori M, Morishita S, Matsushima K. Genome-wide profiling of DNA methylation in human cancer cells. *Genomics*. 98(4):280-7 (2011)

Higasa, K., Nikaido, M., Saito, T.L., Yoshimura, J., Suzuki, Y., Suzuki, H., Nishihara, H., Aibara, M., Ngatunga, B.P., Kalombo, H.W.J., Sugano, S., Morishita, S., Okada, N., Extremely slow rate of evolution in the HOX cluster revealed by comparison between Tanzanian and Indonesian coelacanths. *Gene* 505(2), 2012, 324-332

Mitsui, J., Matsukawa, T., Ishiura, H., Higasa, K., Yoshimura, J., Saito, T.L., Ahsan, B., Takahashi, Y., Goto, J., Iwata, A., Niimi, Y., Riku, Y., Goto, Y., Mano, K., Yoshida, M., Morishita, S., Tsuji, S., CSF1R mutations identified in three families with autosomal dominantly inherited leukoencephalopathy. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics* 159 B(8), 2012, 951-957

2.2.3. 研究開発成果の現在の国際的な評価・位置づけ

代表研究者の発表した論文は、BIRD 関連の主要な 15 編の研究論文被引用数は 821 件、うち海外からの引用が 587 件と、国内のみならず広く海外の研究者へも影響を与えていることが分かる。特に、本研究開発課題においてマルチモーダル統合バイオ DB の構築を行

った出芽酵母のトランスクリプトームに関する論文、メダカ及びカイコゲノムに関する論文の引用が多い。また、脊椎動物の祖先型ゲノムの再構成を行い脊椎動物ゲノムの進化を考察した論文は、本課題の主要論文の中で引用回数が最も多かった。

計算機科学分野においても、多様なデータ構造をもつマルチモーダルデータの取り扱いの理論と実装を実現し、最も評価の高い ACM SIGMOD において論文発表している。

また、UTGB Toolkit のためにサブプロジェクトとして開発した `sqlite-jdbc` (SQL データベースを動かすライブラリ) 及び `snappy-java` (データ圧縮ライブラリ) は、前述のとおり、Maven Central (Java プログラムのための世界共通公開レポジトリ) にて公開・配布されている。ダウンロード数はそれぞれ月 2 万を超えており、広く使われていることが分かる。

2.3. 現在の科学技術研究における研究開発成果の活用状況、発展状況のまとめ

本研究開発で構築したマルチモーダル統合バイオ DB は一般的には著作権で保護する対象であるという考えから、知的財産権は取得せず全て無料で公開されており、それぞれの分野の研究者から多くのアクセス・ダウンロードされている状況にある (2.1.3 データベース・ソフトウェア等へのアクセス数/ダウンロード数、visit 数を参照)。

普段プログラム開発を行わない生物系の研究者が、超高速シークエンサーからのデータを整理・活用するための個人用のゲノムブラウザー (パーソナライズド・ゲノムブラウザ) を作成するために開発された UTGB Toolkit は、ゲノムブラウザーのコードの雛型の作成機能や、データベースへの接続、公開用ウェブサーバーへのゲノムブラウザー設置などを含むさまざまな操作を自動化するユーティリティである UTGB Shell (シェル) を実装し、Tomcat ウェブサーバーを直接起動させることで、手元のデスクトップ上で UTGB を作動させ未公開データを含む実験データを即座に次の解析に回し、視覚的に解析結果を得ることを可能にした。

本研究開発では、様々な生物種のマルチモーダル統合バイオ DB の構築のみならず、マルチモーダルデータを分析・活用することで「ヌクレオソーム構造が DNA 変異の周期性を引き起こす要因となる」という新たな生物学的知見を得たが、マルチモーダル統合バイオ DB を利用したバイオ・ライフサイエンス分野での科学的知見を多数発表・展開している。

- ・ 先の、「ヌクレオソーム構造と DNA 変異」については、武田洋幸研究室 (東大理学系) との共同研究として、JST CREST の支援を受け、発生過程における 3 次元クロマチン構造の動態を観測分析するプロジェクトへと発展している。
- ・ Andrew Fire 研究室 (スタンフォード大学) とは線虫ゲノムの再解読で共同研究を行い、C. elegans の転写開始点について論文発表・web site の公開を行い、世界最大級の線虫のポータルサイト WormBase と相互リンクを作成中である。

- ・ 辻省次研究室（東大医学系）とともに新学術領域研究「ゲノム支援」として、次世代シーケンサー Illumina HiSeq2500 を使いゲノムデータを収集し、家系データ等家系データを使って疾患関連遺伝子を絞り込むことにより、国内の約 30 研究室を支援し、着実に成果を上げている。

上記のいずれの研究も、国際的に十分競争力があり評価されている

3. 研究開発による成果の科学技術的、社会的及び経済的な効果

3.1. バイオインフォマティクス分野及びライフサイエンス分野の研究への貢献

3.1.1. バイオインフォマティクス分野の研究への貢献

本研究では、BIRD プロジェクトが始まって程なく全世界的に急速に普及した超高速シーケンサーに対応し、大規模並列計算手法を確立し超高速シーケンサーのデータ解析を継続してきた。海外でも UCSC ゲノムブラウザーや Ensembl などのゲノムブラウザーや、Galaxy などの環境が提供され、マルチモーダルな解析の基盤が形成されてきている。本研究開発当初は、多様なデータ構造をもつマルチモーダルデータを扱った先行研究例が少なかったなかで、理論と実装方法を構築することで、計算機科学のデータベース分野で最も評価の高い ACM SIGMOD において論文発表している。その後も、関係データベース上の SQL を利用することによるポータビリティの向上、データベースアーカイブ容量を圧縮する方法の研究開発が進展している。

本研究で開発した UTBG Toolkit により、パーソナルな環境で、ゲノムブラウザーを簡単に構築し、超高速シーケンサー等の大量かつ多様なデータの取り扱いの基盤を構築する事が可能となり、様々な研究室で利用されつつある。さらに、本研究開発の成果の論文公表により国際的にも認知され、文献引用および web server へのアクセス・ダウンロードに示されるように、バイオインフォマティクス分野の研究に大きく貢献している。

3.1.2. ライフサイエンス分野（バイオインフォマティクス以外）やその他、科学技術分野の研究への貢献

ゲノム全域から収集されたマルチモーダル生物データを組み合わせることで、新たな科学的発見に至るという研究の動向に貢献している。超高速シーケンサーから得られた大量ゲノムデータを積極的に活用したマルチモーダル統合検索により、ヌクレオソーム構造が DNA 変異の周期性の要因となっていることを示し、新たな生物学的発見が可能であることも実証した。この点の新規性に関しては、以下の記事において取り上げられている。

- ・ Colin A. M. Semple and Martin S. Taylor. The Structure of Change. Science 16 January 2009:347-348.

- ・ Muffato M, Crollius HR. Paleogenomics in vertebrates, or the recovery of lost genomes from the mist of time. *Bioessays*. 30(2):122-34. (2008)
- ・ Sémon M, Wolfe KH. Consequences of genome duplication. *Curr Opin Genet Dev*. 17(6):505-12 (2007)
- ・ Patrick Goymer: Shaping up the genome. *Nature Review*, Vol.7. 79 (2006)

さらに、バイオインフォマティクスに固有なアルゴリズム性能や予測効率の向上に留まらず、科学的発見を見据えてゲノムデータを積極的に活用・展開する事で、辻省次（東大医学系）、武田洋幸（東大理学系）、Andrew Fire（スタンフォード大学）といった生命科学分野の研究者との共同研究が進展しており、ドライとウェットの融合という BIRD 事業の目標も達成・展開されている。

3.2. 人材育成の面から参加研究者の活動状況

3.2.1. 研究開発に参加した研究者のキャリアアップ

本研究開発課題に参加した者で、その後新たなポジションを得た者は6名いた。

3.3. 社会的及び経済的な効果

3.3.1. 研究開発成果が大学や公的研究機関、企業等で、応用に向けて継承または発展した例

代表研究者の研究室出身者が代表取締役役に就任している RNAi 社から、代表研究者との共著で、本研究開発課題で開発された siDirect のアップデート版に関する論文、また東京医科歯科大学等と共著で HPV16 陽性子宮頸がん治療用の E6、E7 siRNA の設計に siDirect を利用した論文が出ており、RNAi 社の製品開発における siRNA 設計に siDirect が利用されている。

RNAi 社の HP には、siDirect に関する解説およびライセンス（非営利学術・研究機関には無償）に関する記載、さらに Enhance siDirect®については、RNAi 社の siRNA 配列設計システムであり、東京大学西郷薫教授、程久美子准教授、森下真一教授等のウェット、ドライ研究成果を基盤に、創薬研究に必須の機能を強化した RNAi 配列設計システムであることが記載されており、RNAi 社において発展している。

3.3.2. その他、研究開発成果が社会的、経済的な効果・効用につながる兆し、可能性

本研究開発課題により構築されたデータベースとソフトウェアは、国内外から多数アクセスされている。また、多数の論文が出版され、ゲノムサイエンス分野にも大きな影響が

あった。また、エピゲノム研究ではセルイノベーションプロジェクトで利用され、パーソナルゲノム研究では2つの新学術領域研究「ゲノム支援」および「パーソナルゲノム情報に基づく脳疾患メカニズムの解明」でも活用されている。

パーソナルゲノム研究において、疾患関連遺伝子発見のための新しいアルゴリズムが生まれれば、医学へ大きく貢献できると考えられる。これまでゲノムと疾患の相関について先駆的研究が展開されてきた脳疾患については、家系もしくは孤発性サンプルから収集されたパーソナルゲノム情報を分析し、疾患と相関の高い遺伝子座を効率的に特定し、配列特異性を適切に特徴化し、背後にある機序を推定する新しい情報学を創成することが可能となっている。実際に代表研究者が、辻省次研究室（東大医学系）とともにしている新学術「ゲノム支援」の成果として発表された下記論文は、超高速 DNA シークエンサーを用い得られた患者のゲノム情報をマルチモーダルな統合バイオ DB により解析することで得られた疾患関連遺伝子についての最新知見である。

Yuji Takahashi, (37 名) Shinichi Morishita, Shoji Tsuji. Mutations in ERBB4 which disrupt the neuregulin-ErbB4 pathway cause amyotrophic lateral sclerosis type 19. *American Journal of Human Genetics, Am J Hum Genet.* 93(5), 900-5, 2013

Isojima T, Doi K, Mitsui J, Oda Y, Tokuhiko E, Yasoda A, Yorifuji T, Horikawa R, Yoshimura J, Ishiura H, Morishita S, Tsuji S, Kitanaka S. A recurrent de novo FAM111A mutation causes Kenny-Caffey syndrome type 2. *J Bone Miner Res.* 2013

Ichikawa Y, Ishiura H, Mitsui J, Takahashi Y, Kobayashi S, Takuma H, Kanazawa I, Doi K, Yoshimura J, Morishita S, Goto J, Tsuji S. Exome analysis reveals a Japanese family with spinocerebellar ataxia, autosomal recessive 1. *J Neurol Sci.* 2013 Aug 15;331(1-2):158-60

Atsushi Ishii, Yoshiaki Saito, Jun Mitsui, Hiroyuki Ishiura, Jun Yoshimura, Hidee Arai, Sumimasa Yamashita, Sadami Kimura, Hirokazu Oguni, Shinichi Morishita, Shoji Tsuji. "Identification of ATP1A3 Mutations by Exome Sequencing as the Cause of Alternating Hemiplegia of Childhood in Japanese Patients." *PLOS ONE* 8.2 (2013)

近年、情報記憶媒体に係るコストは大幅に下がっており、膨大と思われた患者のゲノムデータでさえも、近い将来各大学病院でも十分管理できるようになる可能性は高く、最終的に、個人のゲノムデータはカルテの中に埋め込まれることになると考えられる。医療システムにはゲノム情報だけでなく、多様な表現型情報、病歴、家族歴、投薬履歴といった情報が集約されることを考えると、まさにマルチモーダルな統合バイオ DB となる。マルチモーダル統合バイオ DB からのフィードバックが、新たな医療・施術の基準、有効な薬剤の選択などに利用される可能性は高く、本研究開発課題の成果は医学・医療システムにおいて大きく展開・活用される事が期待される。

資料編

1. 論文リスト

1. #Tomoyuki Yamada, Haruhiko Souma, Shinichi Morishita. PrimerStation: a highly specific multiplex genomic PCR primer design server for the human genome. Nucl. Acids Res. 34: W665-W669 (2006) (査読あり)
2. #*Miura F, Kawaguchi N, Sese J, Toyoda A, Hattori M, Morishita S, and Ito T. A large-scale full-length cDNA analysis to explore the budding yeast transcriptome. Proc Natl Acad Sci U S A. 103(47):17846-51 (2006) (査読あり)
3. Taro L. Saito and Shinichi Morishita. Amoeba Join: Overcoming Structural Fluctuations of XML Data. Proc. of Ninth International Workshop on the Web and Databases (WebDB) (2006). (査読あり)
4. Masahiro Kasahara and Shinichi Morishita. Large-scale genome sequence processing. Imperial College Press, 248pp. (2006) (査読あり)
5. #Masahiro Kasahara(*), Kiyoshi Naruse(*), Shin Sasaki(*), Yoichiro Nakatani(*), Wei Qu, Budrul Ahsan, Tomoyuki Yamada, Yukinobu Nagayasu, Koichiro Doi, Yasuhiro Kasai, Tomoko Jindo, Daisuke Kobayashi, Atsuko Shimada, Atsushi Toyoda, Yoko Kuroki, Asao Fujiyama, Takashi Sasaki, Atsushi Shimizu, Shuichi Asakawa, Nobuyoshi Shimizu, Shin-ichi Hashimoto, Jun Yang, Yongjun Lee, Kouji Matsushima, Sumio Sugano, Mitsuru Sakaizumi, Takanori Narita, Kazuko Ohishi, Shinobu Haga, Fumiko Ohta, Hisayo Nomoto, Keiko Nogata, Tomomi Morishita, Tomoko Endo, Tadasu Shin-I, Hiroyuki Takeda(#), Shinichi Morishita(#), and Yuji Kohara(#). The medaka draft genome and insights into vertebrate genome evolution. Nature 447, 714-719 (2007) * joint first authors. # joint corresponding authors. (査読あり)
6. #Yoichiro Nakatani, Hiroyuki Takeda, Yuji Kohara, and Shinichi Morishita. Reconstruction of the Vertebrate Ancestral Genome Reveals Dynamic Genome Reorganization in Early Vertebrates. Genome Research 17(9): 1254-1265 (2007) (査読あり)
7. #*Budrul Ahsan, Daisuke Kobayashi, Tomoyuki Yamada, Masahiro Kasahara, Shin Sasaki, Taro L. Saito, Yukinobu Nagayasu, Koichiro Doi, Yoichiro Nakatani, Wei Qu, Tomoko Jindo, Atsuko Shimada, Kiyoshi Naruse, Atsushi Toyoda, Yoko Kuroki, Asao Fujiyama, Takashi Sasaki, Atsushi Shimizu, Shuichi Asakawa, Nobuyoshi Shimizu, Shin-ichi Hashimoto, Jun Yang, Yongjun Lee, Kouji Matsushima, Sumio Sugano, Mitsuru Sakaizumi, Takanori Narita, Kazuko Ohishi, Shinobu Haga, Fumiko Ohta, Hisayo Nomoto, Keiko Nogata, Tomomi Morishita, Tomoko Endo, Tadasu Shin-I, Hiroyuki Takeda, Yuji Kohara, and Shinichi

Morishita. UTGB/medaka: genomic resource database for medaka biology. *Nucleic Acids Research*, Vol. 36, Database issue D747-D752 (2008) (査読あり)

8. #Kikuchi Y, Mizuuchi E, Nogami S, Morishita S, Ohya Y. Involvement of Rho-type GTPase in control of cell size in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* 7(4):569-78 (2007) (査読あり)
9. #Ohnuki S, Nogami S, Kanai H, Hirata D, Nakatani Y, Morishita S, Ohya Y. Diversity of Ca²⁺-induced morphology revealed by morphological phenotyping of Ca²⁺-sensitive mutants of *Saccharomyces cerevisiae*. *Eukaryot Cell.* 6(5):817-30. (2007) (査読あり)
10. #Kono K, Nogami S, Abe M, Nishizawa M, Morishita S, Pellman D, Ohya Y. G1/S Cyclin-Dependent Kinase Regulates Small GTPase Rho1p through Phosphorylation of RhoGEF Tus1p in *Saccharomyces cerevisiae*. *Mol Biol Cell.* 19(4):1763-71 (2008) (査読あり)
11. # * The International Silkworm Genome Consortium (S. Morishita as one of corresponding authors). The genome of a lepidopteran model insect, the silkworm *Bombyx mori*. *Insect Biochemistry and Molecular Biology*, Vol. 38, Issue 12, 1036-1045 (2008) (査読あり)
12. # * Shin Sasaki, Cecilia C. Mello, Atsuko Shimada, Yoichiro Nakatani, Shin-ichi Hashimoto, Masako Ogawa, Kouji Matsushima, Sam Guoping Gu, Masahiro Kasahara, Budrul Ahsan, Atsushi Sasaki, Taro Saito, Yutaka Suzuki, Sumio Sugano, Yuji Kohara, Hiroyuki Takeda, Andrew Fire(#), Shinichi Morishita(#) Chromatin-Associated Periodicity in Genetic Variation Downstream of Transcriptional Start Sites. *Science.* 323(5912):401-4 (2009) # joint corresponding authors (査読あり)
13. # * Budrul Ahsan, Taro L. Saito, Shin-ichi Hashimoto, Keigo Muramatsu, Manabu Tsuda, Atsushi Sasaki, Kouji Matsushima, Toshiro Aigaki, and Shinichi Morishita. MachiBase: a *Drosophila melanogaster* 5'-end mRNA transcription database. *Nucleic Acids Research*, Vol. 37, Database issue D49-D53 (2009) (査読あり)
14. Shin-ichi Hashimoto(*), Wei Qu(*), Budrul Ahsan, Katsumi Ogoshi, Atsushi Sasaki, Yoichiro Nakatani, Yongjun Lee, Masako Ogawa, Akio Ametani, Yutaka Suzuki, Sumio Sugano, Clarence C Lee, Robert C Nutter, Shinichi Morishita, Kouji Matsushima. High-resolution analysis of the 5'-end transcriptome using a next generation DNA sequencer. *PLoS One* 4(1):e4108. Epub (2009) (査読あり)
15. # * Taro L. Saito, Jun Yoshimura, Shin Sasaki, Budrul Ahsan, Atsushi Sasaki, Reginaldo Kuroshu and Shinichi Morishita. UTGB Toolkit for Personalized Genome Browsers, *Bioinformatics* 25(15):1856-1861 (2009) (査読あり)
16. # Wei Qu, Shin-ichi Hashimoto, Shinichi Morishita. Efficient frequency-based de novo short read clustering for error trimming in next-generation sequencing.

Genome Research 19(7): 1309-1315 (2009) (査読あり)

17. #Watanabe M, Watanabe D, Nogami S, Morishita S, Ohya Y. Comprehensive and quantitative analysis of yeast deletion mutants defective in apical and isotropic bud growth. *Curr Genet.* 55(4):365-80. (2009) (査読あり)
18. #Naito Y, Yoshimura J, Morishita S, Ui-Tei K. siDirect 2.0: updated software for designing functional siRNA with reduced seed-dependent off-target effect. *BMC Bioinformatics.* 10:392 (2009) (査読あり)
19. Kuroshu RM, Watanabe J, Sugano S, Morishita S, Suzuki Y, Kasahara M. Cost-effective sequencing of full-length cDNA clones powered by a de novo-reference hybrid assembly. *PLoS One.* 7;5(5):e10517 (2010) (査読あり)
20. Hongyan Wu, Taro L Saito, Shinichi Morishita. Accelerating Path-free XML Queries in RDBMS. *IPSJ Online Transactions.* Vol. 3, pp.206-217 (2010)
21. Funakoshi M, Tsuda M, Muramatsu K, Hatsuda H, Morishita S, Aigaki T. A gain-of-function screen identifies wdb and lkb1 as lifespan-extending genes in *Drosophila*. *Biochem Biophys Res Commun.* 405(4):667-72 (2011)
22. Ogoshi K, Hashimoto S, Nakatani Y, Qu W, Oshima K, Tokunaga K, Sugano S, Hattori M, Morishita S, Matsushima K. Genome-wide profiling of DNA methylation in human cancer cells. *Genomics.* 98(4):280-7 (2011)
23. #Qu, W. , Hashimoto, S.-I., Shimada, A., Nakatani, Y., Ichikawa, K., Saito, T.L.a, Ogoshi, K., Matsushima, K., Suzuki, Y., Sugano, S., Takeda, H., Morishita, S., Genome-wide genetic variations are highly correlated with proximal DNA methylation patterns. *Genome Research* 22(8), 2012, 1419-1425
24. Higasa, K., Nikaido, M., Saito, T.L., Yoshimura, J., Suzuki, Y., Suzuki, H., Nishihara, H., Aibara, M., Ngatunga, B.P., Kalombo, H.W.J., Sugano, S., Morishita, S. , Okada, N., Extremely slow rate of evolution in the HOX cluster revealed by comparison between Tanzanian and Indonesian coelacanths. *Gene* 505(2), 2012, 324-332
25. Mitsui, J., Matsukawa, T., Ishiura, H., Higasa, K., Yoshimura, J., Saito, T.L., Ahsan, B., Takahashi, Y., Goto, J., Iwata, A., Niimi, Y., Riku, Y., Goto, Y., Mano, K., Yoshida, M., Morishita, S., Tsuji, S., CSF1R mutations identified in three families with autosomal dominantly inherited leukoencephalopathy. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics* 159 B(8), 2012, 951-957
26. Ishii, A., Saito, Y., Mitsui, J., Ishiura, H., Yoshimura, J., Arai, H., Yamashita, S., Kimura, S., Oguni, H., Morishita, S., Tsuji, S., Sasaki, M., Hirose, S., Identification of ATP1A3 Mutations by Exome Sequencing as the Cause of Alternating Hemiplegia of Childhood in Japanese Patients. *PLoS ONE* 8(2), 2013, Article numbere56120

27. Hashimoto, S.-I. , Ogoshi, K., Sasaki, A., Abe, J., Qu, W., Nakatani, Y.b, Ahsan, B., Oshima, K., Shand, F.H.W., Ametani, A., Suzuki, Y., Kaneko, S., Wada, T., Hattori, M., Sugano, S., Morishita, S., Matsushima, K., Coordinated changes in DNA methylation in antigen-specific memory CD4 T cells. *Journal of Immunology*. 19(8), 2013, 4076-4091
28. Ishikawa, T., Okada, T., Ishikawa-Fujiwara, T., Todo, T., Kamei, Y., Shigenobu, S., Tanaka, M., Saito, T.L., Yoshimura, J., Morishita, S., Toyoda, A., Sakaki, Y., Taniguchi, Y., Takeda, S., Mori, K., ATF6 α/β -mediated adjustment of ER chaperone levels is essential for development of the notochord in medaka fish. *Molecular Biology of the Cell* 24(9), 2013, 1387-1395
29. Saito, T.L., Hashimoto, S.-I., Gu, S.G.cf, Morton, J.J., Stadler, M., Blumenthal, T., Fire, A. , Morishita, S., The transcription start site landscape of *C. elegans*. *Genome Research* 23(8), 2013, 1348-1361
30. Ichikawa, Y., Ishiura, H., Mitsui, J., Takahashi, Y., Kobayashi, S., Takuma, H., Kanazawa, I., Doi, K., Yoshimura, J., Morishita, S., Goto, J., Tsuji, S., Exome analysis reveals a Japanese family with spinocerebellar ataxia, autosomal recessive 1. *Journal of the Neurological Sciences* 331(1), 2013, 158-160

: 本追跡調査において研究代表者が主要な論文として指定したもの（上限 30 報）

* : 研究開発期間終了後の終了報告書において研究代表者が主要な論文として指定したもの

2. 主要論文の被引用回数

論文	国内外別 件数		分野別件数																					出版年別件数								
	国内 件数	海外 件数	農学	生物学・ 生化学	化学	臨床 医学	コンピ ュータ サイエ ンス	経済 学・ 経営 学	工学	環 境・ 生態 学	地球 科学	免疫 学	材料 科学	数学	微生物 学	分子 生物 学・ 遺伝 学	複合 領域	神経 科 学・ 行動 科学	薬理 学・ 毒物 学	物理 学	植物 学・ 動物 学	精神 医 学・ 心理 学	社会 科 学・ 一般	宇宙 科学	2006 年	2007 年	2008 年	2009 年	2010 年	2011 年	2012 年	
1. Tomoyuki Yamada, Haruhiko Souma, Shinichi Morishita. PrimerStation: a highly specific multiplex genomic PCR primer design server for the human genome. Nucl. Acids Res. 34: W665–W669 (2006) (査読あり)	3	5	0	1	1	0	1	0	0	1	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	1	1	4	1	1	0
2. * Miura F, Kawaguchi N, Sese J, Toyoda A, Hattori M, Morishita S, and Ito T. A large-scale full-length cDNA analysis to explore the budding yeast transcriptome. Proc Natl Acad Sci U S A. 103(47):17846–51 (2006)(査読あり)	15	91	0	30	0	6	6	0	0	1	0	0	0	1	5	53	0	0	1	0	2	0	0	0	0	0	14	22	27	11	19	13
5. Masahiro Kasahara(*), Kiyoshi Naruse(*), Shin Sasaki(*), Yoichiro Nakatani(*), Wei Qu, Budrul Ahsan, Tomoyuki Yamada, Yukinobu Nagayasu, Koichiro Doi, Yasuhiro Kasai, Tomoko Jindo, Daisuke Kobayashi, Atsuko Shimada, Atsushi Toyoda, Yoko Kuroki, Asao Fujiyama, Takashi Sasaki, Atsushi Shimizu, Shuichi Asakawa, Nobuyoshi Shimizu, Shin-ichi Hashimoto, Jun Yang, Yongjun Lee, Kouji Matsushima, Sumio Sugano, Mitsuru Sakaizumi, Takanori Narita, Kazuko Ohishi, Shinobu Haga, Fumiko Ohta, Hisayo Nomoto, Keiko Nogata, Tomomi Morishita, Tomoko Endo, Tadasu Shin-I, Hiroyuki Takeda(#), Shinichi Morishita(#), and Yuji Kohara(#). The medaka draft genome and insights into vertebrate genome evolution. Nature 447, 714–719 (2007) * joint first authors. # joint corresponding authors. (査読あり)	111	183	1	82	0	27	3	0	0	5	0	6	0	1	1	129	0	2	0	0	36	0	1	0	—	9	59	68	61	47	50	
6. Yoichiro Nakatani, Hiroyuki Takeda, Yuji Kohara, and Shinichi Morishita. Reconstruction of the Vertebrate Ancestral Genome Reveals Dynamic Genome Reorganization in Early Vertebrates. Genome Research 17(9): 1254–1265 (2007) (査読あり)	11	91	0	39	0	6	1	0	0	0	1	2	0	0	0	43	0	2	0	0	7	0	0	0	—	0	25	24	10	24	19	
7. * Budrul Ahsan, Daisuke Kobayashi, Tomoyuki Yamada, Masahiro Kasahara, Shin Sasaki, Taro L. Saito, Yukinobu Nagayasu, Koichiro Doi, Yoichiro Nakatani, Wei Qu, Tomoko Jindo, Atsuko Shimada, Kiyoshi Naruse, Atsushi Toyoda, Yoko Kuroki, Asao Fujiyama, Takashi Sasaki, Atsushi Shimizu, Shuichi Asakawa, Nobuyoshi Shimizu, Shin-ichi Hashimoto, Jun Yang, Yongjun Lee, Kouji Matsushima, Sumio Sugano, Mitsuru Sakaizumi, Takanori Narita, Kazuko Ohishi, Shinobu Haga, Fumiko Ohta, Hisayo Nomoto, Keiko Nogata, Tomomi Morishita, Tomoko Endo, Tadasu Shin-I, Hiroyuki Takeda, Yuji Kohara, and Shinichi Morishita. UTGB/medaka: genomic resource database for medaka biology. Nucleic Acids Research, Vol. 36, Database issue D747–D752 (2008) (査読あり)	6	7	0	3	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0	3	0	0	0	—	—	2	4	2	4	1	
8. Kikuchi Y, Mizuuchi E, Nogami S, Morishita S, Ohya Y. Involvement of Rho-type GTPase in control of cell size in Saccharomyces cerevisiae. FEMS Yeast Res. 7(4):569–78 (2007) (査読あり)	6	2	0	3	0	1	0	0	0	0	0	0	0	0	1	2	0	0	0	0	1	0	0	0	—	0	0	3	3	2	0	
9. Ohnuki S, Nogami S, Kanai H, Hirata D, Nakatani Y, Morishita S, Ohya Y. Diversity of Ca2+-induced morphology revealed by morphological phenotyping of Ca2+-sensitive mutants of Saccharomyces cerevisiae. Eukaryot Cell. 6(5):817–30. (2007) (査読あり)	6	2	0	2	1	1	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	—	0	0	3	3	1	1	
10. Kono K, Nogami S, Abe M, Nishizawa M, Morishita S, Pellman D, Ohya Y. G1/S Cyclin-Dependent Kinase Regulates Small GTPase Rho1p through Phosphorylation of RhoGEF Tus1p in Saccharomyces cerevisiae. Mol Biol Cell. 19(4):1763–71 (2008)(査読あり)	4	18	0	4	0	2	0	0	0	0	0	0	0	0	3	12	0	0	0	0	1	0	0	0	—	—	1	6	6	3	6	
11. * The International Silkworm Genome Consortium (S. Morishita as one of corresponding authors). The genome of a lepidopteran model insect, the silkworm Bombyx mori. Insect Biochemistry and Molecular Biology, Vol. 38, Issue 12, 1036–1045 (2008)(査読あり)	60	97	2	31	0	18	0	0	1	0	0	0	0	0	5	52	0	3	1	0	43	0	0	0	—	—	0	11	53	47	46	
12. * Shin Sasaki, Cecilia C. Mello, Atsuko Shimada, Yoichiro Nakatani, Shin-ichi Hashimoto, Masako Ogawa, Kouji Matsushima, Sam Guoping Gu, Masahiro Kasahara, Budrul Ahsan, Atsushi Sasaki, Taro Saito, Yutaka Suzuki, Sumio Sugano, Yuji Kohara, Hiroyuki Takeda, Andrew Fire(#), Shinichi Morishita Chromatin-Associated Periodicity in Genetic Variation Downstream of Transcriptional Start Sites. Science. 323(5912):401–4 (2009) # joint corresponding authors (査読あり)	4	44	0	10	0	3	2	0	0	0	0	0	0	0	2	30	0	0	0	0	1	0	0	0	—	—	—	10	19	13	6	

論文	国内外別 件数		分野別件数																				出版年別件数								
	国内 件数	海外 件数	農学	生物学・ 生化学	化学	臨床 医学	コンピ ュータ サイエ ンス	経済 学・ 経営 学	工学	環 境・ 生態 学	地球 科学	免疫 学	材料 科学	数学	微生 物学	分子 生物 学・ 遺伝 学	複合 領域	神経 科学・ 行動 科学	薬理 学・ 毒物 学	物理 学	植物 学・ 動物 学	精神 医学・ 心理 学	社会 科学・ 一般	宇宙 科学	2006 年	2007 年	2008 年	2009 年	2010 年	2011 年	2012 年
15. * Taro L. Saito, Jun Yoshimura, Shin Sasaki, Budrul Ahsan, Atsushi Sasaki, Reginaldo Kuroshu and Shinichi Morishita. UTGB Toolkit for Personalized Genome Browsers, Bioinformatics 25(15):1856-1861 (2009) (査読あり)	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	—	—	—	0	1	0	0
16. Wei Qu, Shin-ichi Hashimoto, Shinichi Morishita. Efficient frequency-based de novo short read clustering for error trimming in next-generation sequencing. Genome Research 19(7): 1309-1315 (2009) (査読あり)	3	13	0	2	0	2	7	0	0	0	0	0	0	0	1	4	0	0	0	0	0	0	0	0	—	—	—	1	4	8	3
17. Watanabe M, Watanabe D, Nogami S, Morishita S, Ohya Y. Comprehensive and quantitative analysis of yeast deletion mutants defective in apical and isotropic bud growth. Curr Genet. 55(4):365-80. (2009) (査読あり)	2	10	1	2	0	1	0	0	0	0	0	0	0	0	2	4	0	0	0	0	2	0	0	0	—	—	—	0	2	7	3
18. Naito Y, Yoshimura J, Morishita S, Ui-Tei K. siDirect 2.0: updated software for designing functional siRNA with reduced seed-dependent off-target effect. BMC Bioinformatics. 10:392 (2009) (査読あり)	3	12	0	4	0	6	1	0	0	0	0	0	0	0	0	1	0	1	2	0	0	0	0	0	—	—	—	0	1	9	5
23. Qu, W. , Hashimoto, S.-I., Shimada, A., Nakatani, Y., Ichikawa, K., Saito, T.L.a, Ogoshi, K., Matsushima, K., Suzuki, Y., Sugano, S., Takeda, H., Morishita, S., Genome-wide genetic variations are highly correlated with proximal DNA methylation patterns. Genome Research 22(8), 2012, 1419-1425	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	—	—	—	—	—	—	0

- ・本追跡調査において研究代表者が主要な論文として指定したもの（上限 30 報）について、トムソン・ロイター社 Web of Science で調査した。調査対象は、2012 年 12 月まで。被引用情報が取得できたものののみ記載した。
- ・論文番号は「資料編 1. 論文リスト」に対応している。
- ・＊：研究開発期間終了後の終了報告書において研究代表者が主要な論文として指定したもの。
- ・[国内外別件数] は、被引用文献の国内外別件数。被引用文献の著者の所属国のうち JAPAN が一つでもあれば国内としている。国内外の合計が全被引用文献数となる。
- ・[分野別件数] は、被引用文献の分野別件数。被引用文献が論文の場合のみカウントしている。
- ・[出版年別件数] は、被引用文献の出版年別件数。「—」は、データなしを表す。

論文	Web of Science 国内外別件数(2012 年 12 月まで)			Google Scholar	
	国内件数	海外件数	合計	～2014 年 4 月 24 日	～2012 年
1. Tomoyuki Yamada, Haruhiko Souma, Shinichi Morishita. PrimerStation: a highly specific multiplex genomic PCR primer design server for the human genome. Nucl. Acids Res. 34: W665–W669 (2006) (査読あり)	3	5	8	17	15
2. ＊Miura F, Kawaguchi N, Sese J, Toyoda A, Hattori M, Morishita S, and Ito T. A large-scale full-length cDNA analysis to explore the budding yeast transcriptome. Proc Natl Acad Sci U S A. 103(47):17846–51 (2006)(査読あり)	15	91	106	161	142
5. Masahiro Kasahara(*), Kiyoshi Naruse(*), Shin Sasaki(*), Yoichiro Nakatani(*), Wei Qu, Budrul Ahsan, Tomoyuki Yamada, Yukinobu Nagayasu, Koichiro Doi, Yasuhiro Kasai, Tomoko Jindo, Daisuke Kobayashi, Atsuko Shimada, Atsushi Toyoda, Yoko Kuroki, Asao Fujiyama, Takashi Sasaki, Atsushi Shimizu, Shuichi Asakawa, Nobuyoshi Shimizu, Shin-ichi Hashimoto, Jun Yang, Yongjun Lee, Kouji Matsushima, Sumio Sugano, Mitsuru Sakaizumi, Takanori Narita, Kazuko Ohishi, Shinobu Haga, Fumiko Ohta, Hisayo Nomoto, Keiko Nogata, Tomomi Morishita, Tomoko Endo, Tadasu Shin-I, Hiroyuki Takeda(#), Shinichi Morishita(#), and Yuji Kohara(#). The medaka draft genome and insights into vertebrate genome evolution. Nature 447, 714–719 (2007) ＊joint first authors. # joint corresponding authors. (査読あり)	111	183	294	502	419

6. Yoichiro Nakatani, Hiroyuki Takeda, Yuji Kohara, and Shinichi Morishita. Reconstruction of the Vertebrate Ancestral Genome Reveals Dynamic Genome Reorganization in Early Vertebrates. Genome Research 17(9): 1254–1265 (2007) (査読あり)	11	91	102	210	155
7. *Budrul Ahsan, Daisuke Kobayashi, Tomoyuki Yamada, Masahiro Kasahara, Shin Sasaki, Taro L. Saito, Yukinobu Nagayasu, Koichiro Doi, Yoichiro Nakatani, Wei Qu, Tomoko Jindo, Atsuko Shimada, Kiyoshi Naruse, Atsushi Toyoda, Yoko Kuroki, Asao Fujiyama, Takashi Sasaki, Atsushi Shimizu, Shuichi Asakawa, Nobuyoshi Shimizu, Shin-ichi Hashimoto, Jun Yang, Yongjun Lee, Kouji Matsushima, Sumio Sugano, Mitsuru Sakaizumi, Takanori Narita, Kazuko Ohishi, Shinobu Haga, Fumiko Ohta, Hisayo Nomoto, Keiko Nogata, Tomomi Morishita, Tomoko Endo, Tadasu Shin-I, Hiroyuki Takeda, Yuji Kohara, and Shinichi Morishita. UTGB/medaka: genomic resource database for medaka biology. Nucleic Acids Research, Vol. 36, Database issue D747–D752 (2008) (査読あり)	6	7	13	22	18
8. Kikuchi Y, Mizuuchi E, Nogami S, Morishita S, Ohya Y. Involvement of Rho-type GTPase in control of cell size in Saccharomyces cerevisiae. FEMS Yeast Res. 7(4):569–78 (2007) (査読あり)	6	2	8	12	12
9. Ohnuki S, Nogami S, Kanai H, Hirata D, Nakatani Y, Morishita S, Ohya Y. Diversity of Ca2+-induced morphology revealed by morphological phenotyping of Ca2+-sensitive mutants of Saccharomyces cerevisiae. Eukaryot Cell. 6(5):817–30. (2007) (査読あり)	6	2	8	15	9
10. Kono K, Nogami S, Abe M, Nishizawa M, Morishita S, Pellman D, Ohya Y. G1/S Cyclin-Dependent Kinase Regulates Small GTPase Rho1p through Phosphorylation of RhoGEF Tus1p in Saccharomyces cerevisiae. Mol Biol Cell. 19(4):1763–71 (2008)(査読あり)	4	18	22	34	29
11. *The International Silkworm Genome Consortium (S. Morishita as one of corresponding authors). The genome of a lepidopteran model insect, the silkworm Bombyx mori. Insect Biochemistry and Molecular Biology, Vol. 38, Issue 12, 1036–1045 (2008)(査読あり)	60	97	157	287	221
12. *Shin Sasaki, Cecilia C. Mello, Atsuko Shimada, Yoichiro Nakatani, Shin-ichi Hashimoto, Masako Ogawa, Kouji Matsushima, Sam Guoping Gu, Masahiro Kasahara, Budrul Ahsan, Atsushi Sasaki, Taro Saito, Yutaka Suzuki, Sumio Sugano, Yuji Kohara, Hiroyuki Takeda, Andrew Fire(#), Shinichi Morishita Chromatin-Associated Periodicity in Genetic Variation Downstream of Transcriptional Start Sites. Science. 323(5912):401–4 (2009) # joint corresponding authors (査読あり)	4	44	48	72	59
13. *Budrul Ahsan, Taro L. Saito, Shin-ichi Hashimoto, Keigo Muramatsu, Manabu Tsuda, Atsushi Sasaki, Kouji Matsushima, Toshiro Aigaki, and Shinichi Morishita. MachiBase: a Drosophila melanogaster 5'-end mRNA transcription database. Nucleic Acids Research, Vol. 37, Database issue D49–D53 (2009) (査読あり)	—	—	—	23	22
15. *Taro L. Saito, Jun Yoshimura, Shin Sasaki, Budrul Ahsan, Atsushi Sasaki, Reginaldo Kuroshu and Shinichi Morishita. UTGB Toolkit for Personalized Genome Browsers, Bioinformatics 25(15):1856–1861 (2009) (査読あり)	0	1	1	9	7
16. Wei Qu, Shin-ichi Hashimoto, Shinichi Morishita. Efficient frequency-based de novo short read clustering for error trimming in next-generation sequencing. Genome Research 19(7): 1309–1315 (2009) (査読あり)	3	13	16	41	31
17. Watanabe M, Watanabe D, Nogami S, Morishita S, Ohya Y. Comprehensive and quantitative analysis of yeast deletion mutants defective in apical and isotropic bud growth. Curr Genet. 55(4):365–80. (2009) (査読あり)	2	10	12	27	15
18. Naito Y, Yoshimura J, Morishita S, Ui-Tei K. siDirect 2.0: updated software for designing functional siRNA with reduced seed-dependent off-target effect. BMC Bioinformatics. 10:392 (2009) (査読あり)	3	12	15	40	29
23. Qu, W. , Hashimoto, S.-I., Shimada, A., Nakatani, Y., Ichikawa, K., Saito, T.L.a, Ogoshi, K., Matsushima, K., Suzuki, Y., Sugano, S., Takeda, H., Morishita, S., Genome-wide genetic variations are highly correlated with proximal DNA methylation patterns. Genome Research 22(8), 2012, 1419–1425	0	0	0	7	3

・ 情報科学・計算機科学分野における研究開発成果では、会議予稿集での発表に引用されることが多いため、Google Scholar での被引用件数を参考情報として添付した。

・ Google での調査方法

対象論文のタイトルを検索し、当該論文を「引用元」とする件数を取得した。

最新の被引用件数と、それから 2013 年以降の被引用件数を引いたものを 2012 年までとした。ただし、年が明確でないものは引かれていないため、2012 年までの被引用件数が実際よりも多くなっている可能性がある。

・ 論文番号は「資料編 1. 論文リスト」に対応している。

・ *：研究開発期間終了後の終了報告書において研究代表者が主要な論文として指定したもの。

3. 学会招待講演・基調講演

1. Shinichi Morishita. Searching Massive Epigenome Data for Evolutionarily Conserved Sequence Motifs. 1st IEEE International Conference on Computational Advances in Bio and medical Sciences. Orland, Florida, USA. Feb. 3, 2011.
2. Shinichi Morishita. Genetic Variation Associated with Nucleosome Structure and DNA Methylation. Fish Genome Meeting 2011, Sanger Center, Cambridge, UK, March 11, 2011.
3. 森下真一 Searching Massive Epigenome Data for Evolutionarily Conserved Sequence Motifs. 第一回データインテンシブサイエンスワークショップ, 本郷, 2011年3月9日
4. Shinichi Morishita. Genetic Variation Associated with Nucleosome Structure and DNA Methylation. The 3rd Congress of BioSoft (BioSoft-2011), Beijing, China, March 23, 2011.
5. 森下真一 生命情報科学若手の会・第2回研究会 三島 2010年10月9日
6. 森下真一 理工系情報学科・専攻協議会 東北大学 2010年8月2日
7. 森下真一 CBRC2009 産総研 CBRC 2009年12月3日
8. 森下真一 人類遺伝学会 第54回大会 2009年9月26日
9. Shinichi Morishita. Chromatin-associated periodicity in genetic variation downstream of transcriptional start sites. The 8th International Workshop on Advanced Genomics. Tokyo, June 17, 2009
10. Shinichi Morishita. Chromatin-associated periodicity in genetic variation downstream of transcriptional start sites. Third International Workshop on Fish Genomes, Sanger Center, Hinxton, UK, March 13, 2009.
11. 森下真一 第31回日本分子生物学会年会・第81回日本生化学会大会 合同大会シンポジウム「超高速シーケンサーとバイオインフォマティクス」2008年12月10日
12. 森下真一 第54回 NIBB コンファレンス ”モデル生物メダカの新たな発展” 2008年2月29日
13. 森下真一 シンポジウム CBRC2007 2007年12月20日
14. Shinichi Morishita. Vertebrate Genome Evolution Examined by Comparing the Human and Fish Genomes. International Symposium Celebrating Dr. David S. Hogness, Recipient of the 23rd International Prize for Biology. Kyoto, Japan, Nov. 21, 2007
15. 森下真一 第七回 PC クラスタシンポジウム 2007年12月13日

16. 森下真一 第 23 回国際生物学賞受賞者 David Hogness 博士を祝う国際シンポジウム 2007 年 11 月 22 日
17. 森下真一 第 13 回小型魚類研究会 2007 年 9 月 16 日
18. 森下真一 日本進化学会第 9 回京都大会ワークショップ「カイコゲノムの解析から見える昆虫の適応進化」 2007 年 9 月 2 日
19. 森下真一 バイオインフォマティクス夏の学校 2007 年 8 月 6 日
20. Shinichi Morishita. The medaka draft genome: new insights into vertebrate genome evolution .Origin of Vertebrates, Tokyo, Japan, March 28th, 2007.

4. 新聞発表等（著作権の関係により非公開）

5. 特許出願・成立

該当なし

6. 学会賞等の受賞

該当なし

7. グラントの獲得実績

1. ゲノム科学の総合的推進に向けた大規模ゲノム情報生産・高度情報解析支援（小原 雄治）：2010-2014 年
2. クロマチン構造と DNA メチル化が誘導する進化（森下 真一 ゲノム生物学 基礎研究(A)）：2011-2014 年
3. CREST 生命動態の理解と制御のための基盤技術の創出 DNA 3 次元クロマチン動態の理解と予測（武田洋幸）分担者 2013-2019 年

8. 書籍等の執筆実績

1. 森下 真一・阿久津 達也 編 「実験医学増刊 生命研究への応用と開発が進むバイオデータベースとソフトウェア最前線」 2008 年 4 月

9. 総説の執筆実績

1. 森下真一, 中谷洋一郎「比較ゲノムを支える情報学：脊椎動物ゲノム進化を推定するロジック」細胞工学 Vol.25,No.8(2006)
2. 中谷洋一郎, 森下真一「メダカゲノム解析に基づく脊椎動物ゲノム進化の推定」実験医学増刊号 Vol.25No.2(2007)
3. 森下真一, 中谷洋一郎「比較ゲノムを支える情報学：脊椎動物ゲノム進化を推定するロジック」細胞工学別冊「比較ゲノム学から読み解く生命システム」(2007)
4. バドルルアーサン、山田智之、森下真一「ゲノムブラウザー」実験医学増刊「生命研究への応用と開発が進むバイオデータベースとソフトウェア最前線」(2008)
5. 曲薇、バドルルアーサン、森下真一「超高速シーケンサーの原理と情報処理」実験医学増刊「生命研究への応用と開発が進むバイオデータベースとソフトウェア最前線」(2008)
6. 森下真一「次世代シーケンサーを用いたゲノムインフォマティクス」実験医学増刊 Vol.27, No.12(2009)
7. 森下真一「次世代高速シーケンサーの特性と情報処理」蛋白質核酸酵素 Vol.54,No.10(2009)
8. 森下真一「生命情報学-ゲノム情報ビッグバンの進展-」情報処理 Vol.51No.6(2010)

10. 参加研究者の活動状況（個人情報が含まれるため非公開）