

平成21年6月22日現在

研究種目：特定領域研究  
 研究期間：2007～2008  
 課題番号：19021049  
 研究課題名(和文)網羅的mRNA絶対定量のためのパイロプライマーの開発  
 研究課題名(英文)Designing pyro-primer sequences for exhaustive quantitation of mRNAs  
 研究代表者  
 藤渕 航 (FUJIBUCHI WATARU)  
 独立行政法人産業技術総合研究所・生命情報工学研究センター・研究チーム長  
 研究者番号：60273512

研究成果の概要:パイロシーケンシングによる細胞内 mRNA 絶対定量を可能にするためのプライマー配列 (パイロプライマー) を開発するため、(1)「遺伝子」発現定量のための低冗長性を許した polyA 近郊パイロプライマー、および (2)「スプライシングバリエーション」を完全定量するエクソン境界近傍パイロプライマーを現実時間内で設計することを目的としそのための方法論とシステムを開発した。

## 交付額

(金額単位：円)

	直接経費	間接経費	合計
2007年度	2,300,000	0	2,300,000
2008年度	2,300,000	0	2,300,000
年度			
年度			
年度			
総計	4,600,000	0	4,600,000

研究分野：生命情報学

科研費の分科・細目：

キーワード：遺伝子発現定量、プライマー配列、配列決定法

## 1. 研究開始当初の背景

PCR プライマーやマイクロアレイプローブの設計は研究がさかんであるが、これを1細胞でのパイロシーケンシング技術による網羅的 mRNA 測定にまで応用したものは論文を調べた限り、まだ存在していなかった。研究代表者は本研究の3年前からプライマー、プローブ設計の研究を目指して、徐々に準備を進めてきていた。日本バイオインダストリー協会(JBA)で遺伝子情報解析に係る知的基盤整備検討準備委員として、国内マイクロアレイ企業、大学関係者と遺伝子発現データの国内標準化事業を行っており、特に標準プローブの設計を担当していた。

## 2. 研究の目的

1 細胞内 mRNA を網羅的に定量するためのパイロシーケンシングプライマーセットを自動的に提示するシステムの開発を行う。mRNA とプライマーの結合定数の条件やクロスハイブリダイゼーションの起こらない条件を満たし、数理計画問題として解くための定式化を行う。具体的には、(1) 遺伝子 mRNA の 3' 側の polyA 付近から発現量を絶対定量できるプライマー、(2) 完全長 cDNA からスプライシングバリエーションを絶対定量できるエクソン 3' 境界近傍プライマーの設計を行う。

## 3. 研究の方法

本プロジェクトの A04 班のパイロシーケンシ

ング法では、実際に50~100bpシーケンシングを計画しているため、これを最大限に生かした mRNA 定量を行うためのプライマーを開発する。特定する mRNA の配列解像度に応じて、(1)「遺伝子」の低冗長性を許した polyA 近郊パイロプライマー、および(2)「スプライシングバリエーション」を定量するエクソン境界近傍パイロプライマーを設計する。(1)では従来よりも「冗長性を許したことにより、より結合能力の高いプライマー」の設計開発を行う。さらに、(2)では、パイロシーケンシングによってエクソンジャンクション配列をシーケンシングすることができるため、「ジャンクションを網羅的かつ正確に区別できるような全エクソン 3' 端近傍(50~100bp)プライマー」を開発する。これらの設計思想を数理計画問題として解く方法を用いて、プライマー配列の候補を提示するための方法及びプログラムを開発する。

#### 4. 研究成果

##### 【システム構成】

高性能プライマーの設計には、大きく分けて 1) 候補配列の全cDNAに対するユニーク性検索、2) プライマーTm値の計算、3) マルチプレックスプライマーのクロスハイブリ性検出の3段階が必要である。

##### (1) 候補配列のユニーク性高速検索

1についてはField Programmable Gate Array (FPGA) 加速ボードを用いた論理回路を設計し、全遺伝子cDNAのプライマー配列ライメントの計算を約1000倍に高速化して32CPUで1年かかる計算を約10日程度にまで短縮した(図1)。

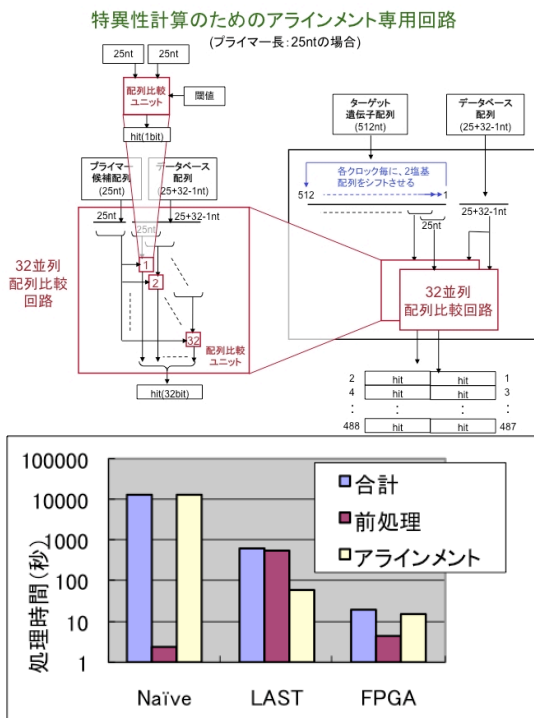


図1: FPGA加速ボードを用いた場合のプライマー配列のユニーク性検索に特化した専用回路(上)とその回路を用いて計算に要する処理時間を単純法(Naive)と高速ソフトウェア(LAST)と比較した結果。前処理を除く時間で比較すると、単純法に比較してLASTで約200倍、FPGAで約1000倍の高速化が達成されている。

本結果については、情報処理学会研究報告会(2008-B10-14)で研究報告を行った。

##### (2) プライマーのTm値の計算

Tm値も数週間で計算できる様に既存のプログラムhybrid-minから高速なバイナリサーチ法を使ってTm値を絞り込むプログラムを作成した。さらに、産総研にあるクラスターマシンによる実装を行った。(図2)

また、非ターゲット遺伝子に対するクロスハイブリダイゼーションを避けるため、ハイブリダイゼーション率を計算し、ターゲット遺伝子に99%結合した場合に5%以上の結合率を有する遺伝子をミス(クロス)ハイブリダイゼーションと定義し、このミスハイブリダイゼーション遺伝子数を与えられた冗長性のパラメーター以下になるように検索するシステムを開発した。

##### プライマーのターゲット遺伝子特性を計算

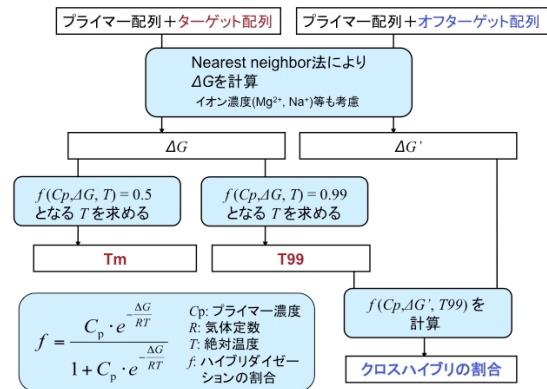


図2: ターゲットと非ターゲットとのクロスハイブリダイゼーションの可能性の計算スキーム

##### (3) マルチプレックスプライマー探索

3は複数のプライマーの条件を同時に探す問題であり、組み合わせ最適化問題に属するため、ギブスサンプリングアルゴリズムを用いた準最適解を高速に解かせる方法を考案した。シーケンサーでは一度に加えることのできるプライマー数は少ない方がよいと考えられるため、与えられたプライマー数の上限値しか用いることができない場合の組み合わせを解くことができるようになった。

さらに、問題を一般化することに成功し、ターゲットにしているK個の遺伝子についてC個以下のプライマーで冗長性を許しながらK個へのユニーク性を最大化(K個に選択されて

いない遺伝子に対するクロスハイブリダイゼーションとの比を最大化) する問題のプロトコルを確立した。(図3)

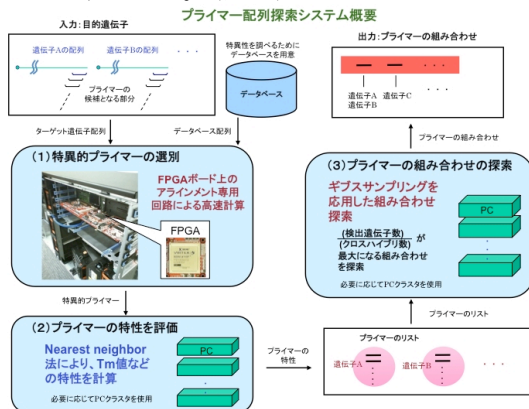


図3：プライマー配列探索システムの計算のフロー概念図

本システムは特許申請準備中である。

【実験による検証】

ヒト乳がん由来 MCF-7 細胞及びマウス初期胚細胞を用いて、発現していると考えられる遺伝子を数個～10個程度選択し、予備検証として設計したプライマーが実際に機能するか PCR 実験による検証を行った。その結果、 $T_m$  値の依存度が大きかったがシステムを用いて配列の  $T_m$  をさらに絞り込み調整することで検出能力の高いプライマーが得られた。

(4) MCF-7 から乳がん遺伝子産物の検出例  
乳がん細胞で発現していると考えられている遺伝子3種について RT-PCR を行ったところ、目的とする mRNA のみが増幅されていることが検証された。この例では、58～68℃程度のアニーリング温度の幅で検出できている。

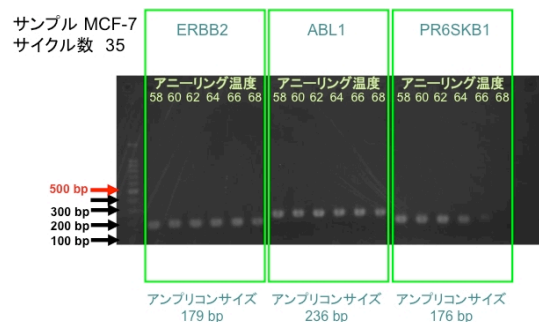


図4：ERBB2, ABL1, PR6SKB1 について PCR を行った結果。目的としているアンプリコンサイズがはっきりとバンドで得られた。

(5) マウス初期胚での転写因子産物の検出例  
マウス初期胚で重要と考えられる転写因子10種を選定し、このプライマーを設計して URR (Universal reference RNA) とマウス初期胚とで検出を比較した。

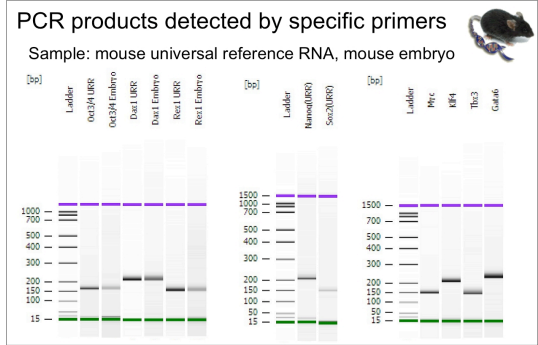


図5：マウスの初期胚からの RNA と URR で同じ転写因子 RNA 産物である Oct3/4, Dax1, Rex1 を増幅した結果 (左)。同じ位置にバンドが得られている。それ以外は URR での結果。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕 (計1件)

Chiba H, Nakagawa S, Taniguchi T, Fujibuchi W. Development of a fast primer search system using FPGA. IPSJ SIG Technical Report, 2008:1-4, 2008. (査読なし)

〔学会発表〕 (計9件)

国際

① Fujibuchi W. “Computer Design of Pyro-Primer Sequences for Exhaustive Quantitation of mRNAs”, The 2nd International Workshop on Approaches to Single-Cell Analysis, 2007/7, Waseda University

② Fujibuchi W. “The Role of Bioinformatics in Single Cell Analysis”, 5<sup>th</sup> International Forum on Post-genome Technologies, 2007/9, Suzhou, China. (invited)

③ Chiba H, Nakagawa S, Taniguchi T, Fujibuchi W. “Computer Design of Pyro-Primer Sequences for Exhaustive Quantitation of mRNAs”, The 3rd International Workshop on Approaches to Single-Cell Analysis, 2008/9, ATH, Zurich

④ Chiba H, Nakagawa S, Taniguchi T, Fujibuchi W. “Designing Pyro-Primer Sequences for Exhaustive Quantitation of mRNAs”, The 7th (International) Life Surveyor Symposium, 2009/1, Hitachi Central Research Laboratory, Tokyo

国内

① 藤渕航「網羅的 mRNA 絶対定量のためのパイロプライマーの開発」、特定領域研究会

- 議、2007年6月、青学会館
- ② 中川草、谷口丈晃、藤渕航「網羅的 mRNA 絶対定量のためのパイロプライマーの開発」、第5回ライフサイエンスシンポジウム、2008年1月、名古屋大学
  - ③ 千葉啓和、中川草、谷口丈晃、藤渕航「網羅的 mRNA 絶対定量のためのパイロプライマーの開発」、第6回ライフサイエンスシンポジウム、2008年6月、大阪大学
  - ④ 千葉啓和、中川草、谷口丈晃、藤渕航「FPGAを用いた高速プライマー配列探索システムの開発」、情報処理学会バイオ情報学研究会、2008年9月、北海道大学
  - ⑤ 千葉啓和、中川草、谷口丈晃、藤渕航「Designing Pyro-primer Sequences for Exhaustive Quantitation of mRNAs」、CBRC2008、2008年11月、産業技術総合研究所・臨海副都心センター別館

[図書] (計0件)

[産業財産権]

○出願状況 (計1件) (準備中)

名称: (仮) ターゲット遺伝子に特異的なマルチプレックスプライマー配列の探索方法、装置、プログラム

発明者: 藤渕航、千葉啓和

権利者: (独) 産業技術総合研究所

種類:

番号:

出願年月日:

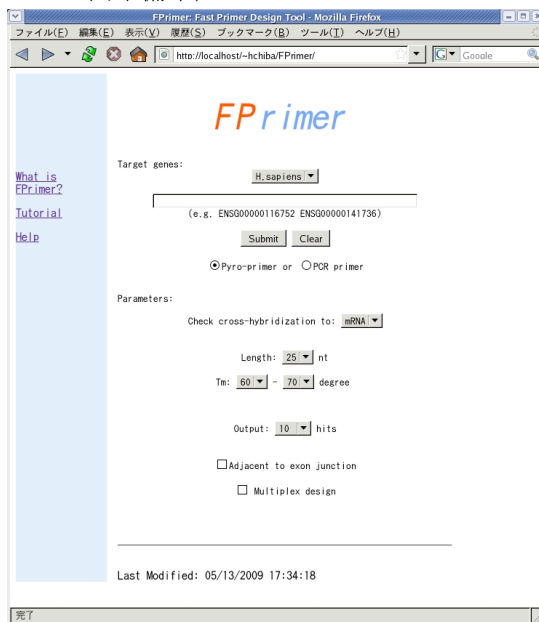
国内外の別: 国内

○取得状況 (計0件)

[その他]

ホームページ等

プライマー探索システムのWebインタラクティブ版のプロトタイプを開発している。(図6: 公開準備中)



#	Position (5-nt)	Primer sequence (5-3')	Tm	Tm	GC%	GC%	GC%	GC%	GC%
1	1-25	TAGCTGTTTCCAAAATATATTGC	55.63	60.90	0.0019	7.46e-06	1.78e-06		
2	2-26	AGCTGTTTCCAAAATATATTGCA	56.15	61.49	0.0088	5.46e-06	1.62e-06		
3	3-27	GCTGTTTCCAAAATATATTGCAA	55.49	60.74	0.0094	6.67e-06	1.65e-06		
4	9-33	TCCAAAATATATTGCAAAATGGAC	54.77	60.06	0.0012	0.00042	2.11e-06		
5	10-34	TCCAAAATATATTGCAAAATGGACA	56.14	61.51	0.0029	0.00028	1.63e-06		
6	11-35	CCAAAATATATTGCAAAATGGACAA	55.02	60.32	0.0075	0.00040	1.36e-06		
7	14-38	AAATATATTGCAAAATGGACAAAGT	54.86	60.56	0.0049	0.00043	9.76e-07		
8	15-39	AATATATTGCAAAATGGACAAAGTG	55.55	61.42	0.0021	0.00024	9.76e-07		
9	16-40	ATATATTGCAAAATGGACAAAGTGG	56.92	62.45	0.0023	0.00018	8.52e-07		
10	17-41	TATATTGCAAAATGGACAAAGTGGG	59.09	64.64	0.0032	8.14e-05	7.21e-07		

#	Position (5-nt)	Primer sequence (5-3')	Tm	Tm	GC%	GC%	GC%	GC%
1	18-42	TATTCACATACATCTAGTTCATCTA	54.63	60.02	0.0218	4.31e-07	6.78e-07	
2	20-44	TTCACATACATCTAGTTCATCTACC	55.50	60.82	0.0124	4.09e-07	5.77e-07	
3	21-45	TACATACATCTAGTTCATCTACCT	56.63	62.05	0.0066	3.85e-07	5.48e-07	
4	22-46	CACATACATCTAGTTCATCTACCTA	55.90	61.49	0.0078	4.01e-07	5.94e-07	
5	23-47	ACATACATCTAGTTCATCTACCTAA	54.51	60.14	0.0717	3.62e-07	6.16e-07	

図6: 遺伝子名(複数)とTm値を入力するとFPGAボードを起動してユニークなプライマー配列セット候補のリストを返すWebシステム「FPrimer」のプロトタイプ。

## 6. 研究組織

### (1) 研究代表者

藤渕 航 (FUJIBUCHI WATARU)

独立行政法人産業技術総合研究所・生命情報工学研究センター・研究チーム長

研究者番号: 60273512

### (2) 研究分担者

なし

### (3) 連携研究者

なし