

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2013 年 4 月 25 日(25.04.2013)



(10) 国際公開番号
WO 2013/058404 A1

- (51) 国際特許分類:
C12N 15/09 (2006.01) *C07K 14/00* (2006.01)
C12Q 1/68 (2006.01)
- (21) 国際出願番号: PCT/JP2012/077274
- (22) 国際出願日: 2012 年 10 月 22 日(22.10.2012)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2011-231346 2011 年 10 月 21 日(21.10.2011) JP
- (71) 出願人: 国立大学法人九州大学(KYUSHU UNIVERSITY, NATIONAL UNIVERSITY CORPORATION) [JP/JP]; 〒8128581 福岡県福岡市東区箱崎六丁目 10 番 1 号 Fukuoka (JP).
- (72) 発明者: 中村 崇裕(NAKAMURA, Takahiro); 〒8128581 福岡県福岡市東区箱崎六丁目 10 番 1 号 国立大学法人九州大学内 Fukuoka (JP). 八木祐介(YAGI, Yusuke); 〒8128581 福岡県福岡市東区箱崎六丁目 10 番 1 号 国立大学法人九州大学内 Fukuoka (JP). 小林 啓子(KOBAYASHI, Keiko); 〒8128581 福岡県福岡市東区箱崎六丁目 10 番 1 号 国立大学法人九州大学内 Fukuoka (JP).
- (74) 代理人: 小野 新次郎, 外(ONO, Shinjiro et al.); 〒1000004 東京都千代田区大手町二丁目 2 番 1 号 新大手町ビル 206 区 ユアサハラ法律特許事務所 Tokyo (JP).
- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), ユーロパ (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:

- 国際調査報告 (条約第 21 条(3))
- 明細書の別個の部分として表した配列リスト (規則 5.2(a))

(54) Title: DESIGN METHOD FOR RNA-BINDING PROTEIN USING PPR MOTIF, AND USE THEREOF

(54) 発明の名称: PPR モチーフを利用した RNA 結合性蛋白質の設計方法及びその利用

(57) Abstract: Provided is a method for designing a protein capable of selectively binding RNA bases or specifically binding an RNA base sequence. This protein includes 1 or more (preferably 2-14) PPR motifs comprising a polypeptide represented by formula (1) and having a length of 30-38 amino acids (in formula (1): Helix A is a moiety which is represented by formula (2), has a length of 12 amino acids, and is capable of forming an α -helix structure (in formula (2), A_1 - A_{12} each independently represent an amino acid); X is a moiety which is either not present or which has a length of 1-9 amino acids; Helix B is a moiety that has a length of 11-13 amino acids, and is capable of forming an α -helix structure; L is a moiety which is represented by formula (3) and has a length of 2-7 amino acids (in formula (3), each amino acid is numbered from the C-terminus in the following manner "i"(-1), "ii"(-2), on condition that in some cases L_{iii} to L_{vii} are not present)). The protein corresponding to the RNA bases or base sequence that will be bound is obtained by combining: A_1 , A_4 , and 3 L_{ii} amino acids; or A_4 and 2 L_{ii} amino acids.

(57) 要約: RNA 塩基選択的に、又は RNA 塩基配列特異的に結合可能な蛋白質を設計する方法を提供する。本発明の蛋白質は、式 1 で表される 30~38 アミノ酸長のポリペプチドからなる PPR モチーフを 1 個以上 (好ましくは 2~14 個) 含む、蛋白質であり (式中: Helix A は、12 アミノ酸長の、 α ヘリックス構造を形成可能な部分であって、式 2 で表され、式 2 中、 A_1 ~ A_{12} はそれぞれ独立にアミノ酸を表し; X は、存在しないか又は 1~9 アミノ酸長からなる部分であり; Helix B は、11~13 アミノ酸長からなる、 α ヘリックス構造を形成可能な部分であり; L は、2~7 アミノ酸長の、式 3 で表される部分であり; 式 3 中、各アミノ酸は、"i"(-1)、"ii"(-2)、と C 末端側からナンバリングされ、ただし、 L_{iii} ~ L_{vii} は存在しない場合がある。) A_1 、 A_4 、及び L_{ii} の 3 つのアミノ酸の組み合わせ、又は A_4 、 L_{ii} の 2 つのアミノ酸の組み合わせを、対象 RNA 塩基又は塩基配列に応じたものとする。

WO 2013/058404 A1

明 細 書

発明の名称：

PPRモチーフを利用したRNA結合性蛋白質の設計方法及びその利用

技術分野

[0001] 本発明は、意図したRNA塩基又はRNA配列に選択的又は特異的に結合可能な蛋白質に関する。本発明では、pentatricopeptide repeat (PPR)モチーフを利用する。本発明は、RNA結合性蛋白質の同定、設計、PPR蛋白質の標的RNAの同定、RNAの機能制御のために用いることができる。本発明は、医療分野、農学分野等で有用である。

背景技術

[0002] 近年、様々な解析より明らかになった核酸結合性の蛋白質因子を用いて、意図する配列に結合する技術が確立、利用されている。この配列特異的な結合を利用することで、標的とする核酸（DNA又はRNA）の細胞内局在の解析、標的とするDNA配列の除去、又はその下流に存在する蛋白質コード遺伝子の発現の制御（活性化、又は不活化）が可能になりつつある。

[0003] DNAに作用する蛋白質性因子として、ジンクフィンガー蛋白質(非特許文献1)やTAL effector(非特許文献2、特許文献1)を蛋白質工学的材料とした研究及び開発が行われているが、RNAに特異的に作用する蛋白質性因子の開発はいまだ非常に限定されている。これは、一般的に蛋白質を構成するアミノ酸配列が有するRNAとの親和性及び結合RNA配列との法則性がほとんど明らかになっていない、もしくは法則性が見いだせないためである。例外的に、38アミノ酸から成るpufモチーフの複数個の繰り返しで構成されるpumilio蛋白質に関して、pufモチーフ1個がRNA 1塩基に結合することが示されており(非特許文献3)、pumilio蛋白質を用いた新規RNA結合特性をもつ蛋白質、及びRNA結合特性の改変技術が試みられている(非特許文献4)。しかし、pufモチーフは高度に保存されており、かつ存在数が少ない。そのため、限られたRNA配列に作用する蛋白質因子の創成にしか用いられていない。

- [0004] 他方、ゲノム配列情報から、植物のみで500個もの大きなファミリーを形成する蛋白質、PPR蛋白質（pentatricopeptide repeat (PPR)モチーフを有するタンパク質）が同定された（非特許文献5）。PPR蛋白質は核コードであるが、専らオルガネラ（葉緑体とミトコンドリア）のRNAレベルでの制御、切断、翻訳、スプライシング、RNA編集、RNA安定性に遺伝子特異的に作用する。PPR蛋白質は、典型的には、保存性の低い35アミノ酸のモチーフ、すなわちPPRモチーフが約10個連続した構造を有しており、PPRモチーフの組み合わせが、RNAとの配列選択的な結合を担っていると考えられている。ほとんどのPPR蛋白質はPPRモチーフ約10個の繰り返しのみで構成されており、多くの場合、触媒作用を発揮するために必要なドメインが見いだせない。そのため、このPPR蛋白質の実体はRNAアダプターだと考えられている（非特許文献6）。
- [0005] 本発明者らは、このPPRモチーフを利用した、RNA結合性蛋白質の改変方法について、提案してきた（特許文献2）。

先行技術文献

特許文献

- [0006] 特許文献1：W02011/072246
特許文献2：W02011/111829

非特許文献

- [0007] 非特許文献1：Maeder, M.L., Thibodeau-Beganny, S., Osiak, A., Wright, D.A., Anthony, R.M., Eichinger, M., Jiang, T., Foley, J.E., Winfrey, R.J., Townsend, J.A., et al. (2008). Rapid "open-source" engineering of customized zinc-fingernucleases for highly efficient gene modification. *Mol. Cell* 31, 294-301.
- 非特許文献2：Miller, J.C., Tan, S., Qiao, G., Barlow, K.A., Wang, J., Xia, D.F., Meng, X., Paschon, D.E., Leung, E., Hinkley, S.J., et al. (2011). A TALE nuclease architecture for efficient genome editing. *Nature biotech.* 29, 143-148.
- 非特許文献3：Wang, X., McLachlan, J., Zamore, P.D., and Hall, T.M. (2

002). Modular recognition of RNA by a human pumilio-homology domain. Cell 110, 501-512.

非特許文献4: Cheong, C.G., and Hall, T.M. (2006). Engineering RNA sequence specificity of Pumilio repeats. Proc. Natl. Acad. Sci. USA 103, 13635-13639.

非特許文献5: Small, I.D., and Peeters, N. (2000). The PPR motif - a TPR-related motif prevalent in plant organellar proteins. Trends Biochem. Sci. 25, 46-47.

非特許文献6: Woodson, J.D., and Chory, J. (2008). Coordination of gene expression between organellar and nuclear genomes. Nature Rev. Genet. 9, 383-395.

発明の概要

発明が解決しようとする課題

[0008] PPR蛋白質のRNAアダプターとしての性質は、PPR蛋白質を構成するそれぞれのPPRモチーフの性質、及び複数個のPPRモチーフの組み合わせで決定されると予想される。しかしながら、そのアミノ酸構成と機能の相関関係はほとんど明らかでない。PPRモチーフがRNA結合特性を発揮する際に機能するアミノ酸を同定し、またPPRモチーフの構造と標的塩基との関係が明らかになれば、PPRモチーフの構造や複数のモチーフの組み合わせを人為的に操作することで、任意の配列、長さを有するRNAに結合可能な蛋白質を構築できる可能性がある。

課題を解決するための手段

[0009] 本発明者らは、上記課題を解決するために、遺伝学的に解析されたPPR蛋白質、特にRNA編集（RNAレベルでの遺伝情報の改変、特にシトシン(以下C)からウラシル(以下U)への変換）に関わるPPR蛋白質とその標的RNA配列について検討してきた。そして、計算科学的手法を用いることで、PPRモチーフ中の、3つのアミノ酸（1番、4番、“ii”（-2）番のアミノ酸）に、特定のRNA塩基との結合を司る情報が含まれていることを明らかにした。より詳細には、本発

明者らは、PPRモチーフの結合RNA塩基の選択性（特異性ということもある。）は、モチーフを構成する2つの α ヘリックス構造のうち、最初のヘリックスに含まれる1番目及び4番目のアミノ酸、及び2つ目のヘリックスの後の、ループ構造を形成しうる部分において、後ろ（C末端側）から2番目のアミノ酸（“ii”；-2番）、の3箇所のアミノ酸で決定されることを見出し、本発明を完成した。

[0010] 本発明は、以下を提供する：

[1] RNA塩基選択的に、又はRNA塩基配列特異的に結合可能な蛋白質を設計する方法であって；

蛋白質が、式1で表される30～38アミノ酸長のポリペプチドからなるPPRモチーフを1個以上（好ましくは2～14個）含む、蛋白質であり

[0011] [化1]

(Helix A)-X-(Helix B)-L (式1)

（式中：

Helix Aは、12アミノ酸長の、 α ヘリックス構造を形成可能な部分であって、式2で表され、

[0012] [化2]

$A_1-A_2-A_3-A_4-A_5-A_6-A_7-A_8-A_9-A_{10}-A_{11}-A_{12}$ (式2)

式2中、 $A_1 \sim A_{12}$ はそれぞれ独立にアミノ酸を表し；

Xは、存在しないか又は1～9アミノ酸長からなる部分であり；

Helix Bは、11～13アミノ酸長からなる、 α ヘリックス構造を形成可能な部分であり；

Lは、2～7アミノ酸長の、式3で表される部分であり；

[0013]

[化3]



式3中、各アミノ酸は、“i” (-1)、“ii” (-2)、とC末端側からナンバリングされ、

ただし、 $L_{iii} \sim L_{vii}$ は存在しない場合がある。)

A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせ、又は A_4 、 L_{ii} の2つのアミノ酸の組み合わせを、対象RNA塩基又は塩基配列に応じたものとする、方法。

[2] A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせを、RNA塩基又は対象塩基配列に応じたものとする、[1]に記載の方法であって、アミノ酸の組み合わせが、下記のいずれかに基づいて決定される、方法：

(3-1) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸が、順に、バリン、アスパラギン、アスパラギン酸であるとき、そのPPRモチーフは、U（ウラシル）に選択的に結合でき；

(3-2) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸が、順に、バリン、トレオニン、アスパラギンであるとき、そのPPRモチーフは、A（アデニン）に選択的に結合でき；

(3-3) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸が、順に、バリン、アスパラギン、アスパラギン酸であるとき、そのPPRモチーフは、C（シトシン）に選択的に結合でき；

(3-4) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸が、順に、グルタミン酸、グリシン、アスパラギン酸であるとき、そのPPRモチーフは、G（グアニン）に選択的に結合でき；

(3-5) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸が、順に、イソロイシン、アスパラギン、アスパラギン酸であるとき、そのPPRモチーフは、C又はUに選択的に結合でき；

(3-6) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸が、順に、バリン、トレオニン、ア

スパラギン酸であるとき、そのPPRモチーフは、Gに選択的に結合でき；

(3-7) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸が、順に、リジン、トレオニン、アスパラギン酸であるとき、そのPPRモチーフは、Gに選択的に結合でき；

(3-8) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸が、順に、フェニルアラニン、セリン、アスパラギンであるとき、そのPPRモチーフは、Aに選択的に結合でき；

(3-9) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸が、順に、バリン、アスパラギン、セリン、の場合であるとき、そのPPRモチーフは、Cに選択的に結合でき；

(3-10) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸が、順に、フェニルアラニン、トレオニン、アスパラギンであるとき、そのPPRモチーフは、Aに選択的に結合でき；

(3-11) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸が、順に、イソロイシン、アスパラギン、アスパラギン酸であるとき、そのPPRモチーフは、UまたはAに選択的に結合でき；

(3-12) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸が、順に、トレオニン、トレオニン、アスパラギンであるとき、そのPPRモチーフは、Aに選択的に結合でき；

(3-13) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸が、順に、イソロイシン、メチオニン、アスパラギン酸であるとき、そのPPRモチーフは、UまたはCに選択的に結合でき；

(3-14) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸が、順に、フェニルアラニン、プロリン、アスパラギン酸であるとき、そのPPRモチーフは、Uに選択的に結合でき；

(3-15) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸が、順に、チロシン、プロリン、アスパラギン酸であるとき、そのPPRモチーフは、Uに選択的に結合でき；

(3-16) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸が、順に、ロイシン、トレオニン、アスパラギン酸であるとき、そのPPRモチーフは、Gに選択的に結合できる。

[3] A_4 、及び L_{ii} の2つのアミノ酸の組み合わせを、RNA塩基又は対象塩基配列に応じたものとする、[1]に記載の方法であって、アミノ酸の組み合わせが、

下記のいずれかに基づいて決定される、方法：

(2-1) A_4 、 L_{ii} が、順に、アスパラギン、アスパラギン酸であるとき、そのモチーフは、Uに選択的に結合でき；

(2-2) A_4 、 L_{ii} が、順に、アスパラギン、アスパラギンであるとき、そのモチーフは、Cに選択的に結合でき；

(2-3) A_4 、 L_{ii} が、順に、トレオニン、アスパラギンであるとき、そのモチーフは、Aに選択的に結合でき；

(2-4) A_4 、 L_{ii} が、順に、トレオニン、アスパラギン酸であるとき、そのモチーフは、Gに選択的に結合でき；

(2-5) A_4 、 L_{ii} が、順に、セリン、アスパラギンであるとき、そのモチーフは、Aに選択的に結合でき；

(2-6) A_4 、 L_{ii} が、順に、グリシン、アスパラギン酸であるとき、そのモチーフは、Gに選択的に結合でき；

(2-7) A_4 、 L_{ii} が、順に、アスパラギン、セリンであるとき、そのモチーフは、Cに選択的に結合でき；

(2-8) A_4 、 L_{ii} が、順に、プロリン、アスパラギン酸であるとき、そのモチーフは、Uに選択的に結合でき；

(2-9) A_4 、 L_{ii} が、順に、グリシン、アスパラギンであるとき、そのモチーフは、Aに選択的に結合でき；

(2-10) A_4 、 L_{ii} が、順に、メチオニン、アスパラギン酸であるとき、そのモチーフは、Uに選択的に結合でき；

(2-11) A_4 、 L_{ii} が、順に、ロイシン、アスパラギン酸であるとき、そのモチーフは、Cに選択的に結合でき；

(2-12) A_4 、 L_{ii} が、順に、バリン、トレオニンであるとき、そのモチーフは、Uに選択的に結合できる。

[4] [1]に定義されたPPRモチーフを、1個以上（好ましくは2～14個）含む、RNA結合性蛋白質の標的となる塩基、又は塩基配列を同定する方法であって：

同定が、2に記載の(3-1)～(3-16)のいずれか、又は3に記載の(2-1)～(2-12

)のいずれかに基づいて、PPRモチーフの A_1 、 A_4 、及び L_{ij} の3つのアミノ酸の組み合わせ、又は A_4 、 L_{ij} の2つのアミノ酸の組み合わせに応じた塩基の有無を認定することにより行われる、方法。

[5] 標的RNA塩基、又は特定の塩基配列を有する標的RNAに結合可能な、[1]に定義されたPPRモチーフを1個以上（好ましくは2～14個）含む、PPR蛋白質を同定する方法であって：

同定が、2に記載の(3-1)～(3-16)のいずれか、又は3に記載の(2-1)～(2-12)のいずれかに基づいて、標的RNA塩基、又は標的RNAを構成する特定の塩基に応じた、PPRモチーフの A_1 、 A_4 、及び L_{ij} の3つのアミノ酸の組み合わせの有無を認定することにより行われる、方法。

[6] [1]に記載された方法で設計された蛋白質を用いる、RNAの機能の制御方法。

[7] [1]に記載された方法で設計された蛋白質からなる領域と機能性領域とが連結されてなる、複合体。

[8] 以下の工程を含む、細胞の遺伝物質を改変する方法：

標的配列を有するRNAを含む細胞を準備し；そして

[7]に記載された複合体を細胞に導入することにより、複合体の蛋白質領域が標的配列を有するRNAに結合し、そのため機能性領域が、標的配列を有するRNAを改変する、方法。

[9] 細胞質雄性不稔性の稔性回復因子として働くPPR蛋白質遺伝子において、様々な品種間で見られる当該遺伝子のアミノ酸多型を検出する工程；

当該遺伝子における多型と稔性との関連を特定する工程；

被検サンプルから得られたPPR蛋白質遺伝子の塩基配列を特定し、被検サンプルの稔性を決定する工程；

を含む、PPR蛋白質遺伝子の稔性を判定する方法。

[10] PPR蛋白質が、[1]において定義される式1で表される30～38アミノ酸長のポリペプチドからなるPPRモチーフを1個以上（好ましくは2～16個）含む、蛋白質である、[9]に記載の方法。

[11] アミノ多型をPPRモチーフごとの多型として特定することを特徴とする、[9]または[10]に記載の方法。

[12] PPRモチーフの多型が、式1のモチーフの A_1 、 A_4 、及び L_{ij} の3つのアミノ酸の組み合わせ、又は A_4 、 L_{ij} の2つのアミノ酸の組み合わせにより特定されるものである、[9]～[11]のいずれかに記載の方法。

[13] PPRモチーフの多型が、式1のモチーフの4番アミノ酸 (A_4) の多型により特定される、請求項12に記載の方法。

[14] PPR蛋白質遺伝子上のすべてのPPRモチーフの4番アミノ酸が、Enko Bの対応するすべてのPPRモチーフの4番アミノ酸と同一であることが、稔性であることを示す、請求項13に記載の方法。

[15] PPR蛋白質遺伝子が、orf687様遺伝子（すなわち、Enko Bをコードする「687遺伝子」と相同な遺伝子座に座乗するファミリー遺伝子、Enko Bと90%以上のアミノ酸配列同一性を有する遺伝子、Enko Bをコードする「ORF687遺伝子」90%以上の塩基配列同一性を有する遺伝子）である、請求項9～14のいずれか1項に記載の方法。

[16] 様々な品種のorf687様遺伝子によりコードされる蛋白質が、配列番号：576～578、585～591のいずれかである、請求項9～15のいずれか1項に記載の方法。

発明の効果

[0014] 本発明により、対象RNA塩基に結合可能なPPRモチーフ及びそれを含む蛋白質が提供できる。複数個のPPRモチーフを配することにより、任意の配列や長さを有する標的RNAに結合可能な蛋白質が提供できる。

[0015] 本発明により、任意のPPR蛋白質の標的RNAを予測し、同定することができ、また逆に、任意のRNAに結合するPPR蛋白質を予測し、同定することができる。標的RNA配列を予測することで、その遺伝子的実体を明らかにし、また利用できる可能性が広がる。例えば、本発明においてPPR蛋白質の機能として稔性を考える場合、細胞質雄性不稔性の回復因子として働くような産業的に有用なPPR蛋白質遺伝子について、様々なアミノ酸多型を有した相同遺伝子の機

能性をその標的RNA配列の相違から検定することができる。

[0016] また、本発明により提供されるPPRモチーフ又はPPR蛋白質に機能性を結合し、複合体を調製することができる。

[0017] さらに本発明により、上記の複合体を生体内にデリバリーし、機能させる方法、又は本発明により得られた蛋白質をコードする核酸配列（DNA、RNA）を用いた形質転換体の作製や、生物（細胞、組織、個体）における様々な場面での、特異的な改変、制御及び機能の付与に利用できる。

図面の簡単な説明

[0018] [図1]図1は、PPRモチーフの保存配列とアミノ酸番号を示す。(A) 本発明で定義するPPRモチーフを構成するアミノ酸、およびそのアミノ酸番号を記す。(B) 結合塩基選択性を制御する3つのアミノ酸（1、4、“ii”（-2）番）の予測構造上での位置を示す。(C) 予測構造上での当該アミノ酸の位置。シロイヌナズナCRR4（配列番号6）、CRR21（配列番号3）の全アミノ酸を問い合わせ配列に用いて、PHYRE (<http://www.sbg.bio.ic.ac.uk/phyre/>) プログラムで予測構造を解析したところ、O - GlucNAc transferase(1w3b)を鋳型に、それぞれ高スコアで構造が予測された（ $4.3e-17$ and $4.7e-16$; for CRR4 and CRR21）。そのうち、CRR4の5番目のPPRモチーフ（左図）およびCRR21の8番目のPPRモチーフ（右図）を示した。1、4、“ii”（-2）番はマゼンタ色（白黒表示では濃灰色）のスティックとして示した。

[図2]図2は、これまでに解析されたRNA編集PPR蛋白質とその標的となるRNA編集部位を示す。

[図3-1]図3は、シロイヌナズナRNA編集PPR蛋白質のPPRモチーフ配列とアミノ酸番号を示す。

[図3-2]図3-2は、図3-1の続きを示す。

[図3-3]図3-3は、図3-2の続きを示す。

[図3-4]図3-4は、図3-3の続きを示す。

[図4]図4は、RNA認識に関わるPPRモチーフ中のアミノ酸を示す。(A) PPRモチーフ中の結合塩基指定能力を持つアミノ酸の同定。RNA編集PPR蛋白質のPPRモ

モチーフをRNA編集部位上流配列と様々な配置で整列させた。整列は、モチーフと塩基を1対1として、連続的に配置することで行った。整列P1は最後のPPRモチーフを編集されるCの一つ前の塩基に対応させた。塩基配列を右方向に1塩基ずつずらすことでP2～P6の整列を得た。四角はPPRモチーフ、菱形はC末端側の付加モチーフ（E、E⁺、DYW）を示す。もしモチーフ内の特定箇所のアミノ酸（例、緑色（白黒表示では濃灰色）のモチーフのアミノ酸）がRNA塩基認識を担う場合、特定の整列に置いて、対応する塩基との間にlow randomnessが期待できる（右下図）。そうでない場合、high randomnessが予想される（右上図）。(B) 1、4、“ii”（-2）番アミノ酸の結合RNA塩基指定能力。各整列におけるアミノ酸と塩基のlow randomnessをP値で示した。(C) 様々な核酸の分類における1、4、“ii”（-2）番アミノ酸の結合RNA塩基指定能力。(B)と同じ。ただし、核酸はプリンまたはピリミジン（RY; A&G or U&C）、水素結合グループ（WS; A&U、G&C）の分類。(D) 上記(C)で示したPPRモチーフ中のRNA認識アミノ酸による結合塩基指定能力についてさらに詳細に解析した結果、4番アミノ酸が結合する塩基のプリン/ピリミジン（RY）を主に区別する以外に、“ii”（-2）番アミノ酸が塩基のアミノ型（AおよびC）ケト型（GおよびU）（MK）の区別に働くことが示された（図4D）。(E) いくつかのPPRモチーフのRNA認識コード（PPRコード）の例。白地は1、4、“ii”（-2）番アミノ酸の種類。各コードの出現頻度をNo.に、また対応する核酸の出現頻度をNucleotide frequencyに示した。

[図5]図5は、RNA認識に関わるPPRモチーフ中のアミノ酸の同定(例)を示す。各整列におけるPPRモチーフと対応するRNA塩基のデータセットを用いて、RNA認識に関わるアミノ酸を探索した。例えば、整列P4におけるPPRモチーフと対応するRNA塩基のデータを用い、4番および5番アミノ酸の結合RNA塩基指定能力を解析した。それぞれの整列において、まずアミノ酸の種類でデータを並べ替え、含まれるRNA塩基の数を算出した（左上図）。つぎに理論値をデータセットに含まれる全てのRNAの出現頻度の中央値を基に作製した（右上図）。この2つのデータを用いたカイ二乗検定により、P値を算出した。上図は、有

意なP値が得られた整列P4の4番アミノ酸、下図は有意なP値が得られなかった整列4の5番アミノ酸、における解析結果を示す。

[図6]図6は、RNA塩基指定能力を担うアミノ酸の探索を示す。(A) アミノ酸の種類と塩基の出現頻度との間のlow randomnessをP1～P6の整列における全ての位置のアミノ酸に対して算出した。有意なP値 ($P < 0.01$)を示したアミノ酸は、マゼンタ色（白黒表示では、濃灰色）で示した。シアン色の線（グラフ中の横線。白黒表示では濃灰色）は、 P 値=0.01を示す。(B) 各整列におけるlow randomnessのまとめ。(A)で示した各位置のアミノ酸でのP値の積をその整列におけるlow randomnessの総合値として示した。

[図7]図7は、2つのアミノ酸による結合RNA塩基指定能力を示す。異なる組み合わせの2アミノ酸（1&4、1&“ii”、4&“ii”番アミノ酸）による結合RNA塩基指定能力を図4と同様に、アミノ酸と対応する塩基とのlow randomnessで解析した。

[図8]図8は、シロイヌナズナより抽出したPPRモチーフのRNA認識コードを示す。

[図9]図9は、ヒメツリガネゴケRNA編集PPR蛋白質、及び各蛋白質が働くRNA編集部位の配列を示す。蛋白質のモチーフ構造と共に、各PPRモチーフ中の1、4、“ii”（-2）番アミノ酸配列を示した。マゼンタ、シアン色（いずれも白黒表示では濃灰色）の文字は、それぞれシロイヌナズナより抽出されたtriPPR、またはdiPPRコードと相同なアミノ酸の組み合わせを示す。C末端側の付加モチーフ（E、E+、DYW）も示した。各蛋白質が作用するRNA編集部位の配列（編集されるCを含む上流配列）を図4で示した整列P4の位置で示した。

[図10]図10は、PPR蛋白質とRNA編集部位RNA配列との適合値算出法のフローチャートを示す。Uniprot、もしくはPROSITEデータベースより、蛋白質のPPRモデルを入手し、図1に従って、各アミノ酸番号を付与する。1、4、“ii”番アミノ酸を抽出する。例として、コケPPR蛋白質、PpPPR71を示した。次に一致するアミノ酸の組み合わせを、triPPRコード行列に変換する。triPPRコードに変換できなかったモチーフは、次にdiPPRコード行列に変換する。平行して

、RNA編集部位30nt（最後が編集されるC）を数式行列に変換する。例として、PpPPR71蛋白質が作用するccmFCeU122SF配列を示した。次に、蛋白質コード行列とRNA数式行列の相当するマス目同士の積を求め、その和から適合値を算出する。蛋白質コード行列の最後の行は編集されるCの4つ前の塩基に相当する行に適合させること。この計算は、triPPRコード、diPPRコードから作製した蛋白質コード行列それぞれで行う。複数のRNA配列との適合値より作製した標準分布曲線を用い、それぞれのRNA配列に対する暫定P値をtriPPRコード、diPPRコード、それぞれで求める。最終的な適合値（P値）は、triPPRおよびdiPPRコードの暫定P値の積として算出する。

[図11]図11は、PPRコードを用いたPPR蛋白質の標的RNA配列の予測を示す。(A) コケPPR蛋白質より、1、4、“ii”（-2）番アミノ酸を抽出し、図10に示したようにtriPPRまたはdiPPRコードに変換し、RNA編集部位との適合値を計算、P値で示した。RNA編集部位として、コケの13箇所のRNA編集部位、参照配列として、シロイヌナズナ葉緑体の34 RNA編集部位を用いた。図には、コケの13 RNA編集部位との適合値のみを示した。菱形は、それぞれの編集部位に対する各蛋白質の適合値を示す。正しい編集部位をマゼンタ色（白黒表示ではベタ灰色）で示した。(B) (A)に記したP値をテーブルで示した。

[図12]図12は、シロイヌナズナRNA編集蛋白質を用いたRNA編集部位予測精度の検証を示す。コード抽出に用いたシロイヌナズナPPR蛋白質を用いて、予測精度を検証した。(A) 葉緑体RNA編集部位全34箇所に対する既知PR蛋白質13コのRNA編集部位予測。それぞれの菱形は、蛋白質とRNA編集部位配列との適合値を示す。正しいRNA編集部位をマゼンタ色（白黒表示ではベタ灰色）で示した。(B) ミトコンドリアRNA編集部位全488箇所に対する既知PR蛋白質11コのRNA編集部位予測。

[図13]図13は、シロイヌナズナPPR蛋白質AHG11の標的RNA編集部位の予測と実験的な検証を示す。(A) AHG11のモチーフ構造。12コのPPRモチーフとC末端側の付加モチーフ（E、E+、DYW）という典型的なRNA編集型PPR蛋白質の構造を有する。Ahg11変異体では、星印で示した位置（295番目のTrp）の点変異によ

リコード領域内に新たな翻訳集結コドンが見いだせる。(B) シロイヌナズナの葉緑体、ミトコンドリアに含まれる全てのRNA編集部位を用いた標的RNA編集部位予測。最も高いP値を示したトップ10編集部位を示した。野生株、変異株におけるRNA編集の有無を実験的に検証し、editing statusとして示した。野生株、変異株両方で編集が検出された部位をE、変異株のみでRNA編集が見いだせなかった部位をUn、と示した。(C) 予測の結果をグラフで示した。(D)

AHG11の標的RNA編集部位の実験的な検証。ミトコンドリアnad4を含む領域のシーケンス解析結果を示す。野生株、ahg11変異株よりRNAを抽出し、逆転写によって、cDNAを調製し、塩基配列解析を行った。この領域には、2つのRNA編集部位 (nsd4_362及び_376) が存在する。編集された部位を黒矢印、編集されていない部位を白矢印、で示した。

[図14]図14は、葉緑体ゲノム配列からの標的部位の予測を示す。6コのPPR蛋白質を用いて、シロイヌナズナ葉緑体全ゲノム配列 (154,478bp) から標的部位の予測を行った。予測には、シロイヌナズナより抽出したコード (At code)、もしくは、シロイヌナズナとコケから抽出したコード (At+Pp code) を用いた。

[図15]図15は、シロイヌナズナ及びヒメツリガネゴケより抽出したPPRモチーフのRNA認識コードを示す。

[図16-1]図16は、本発明に関連するアミノ酸配列又は塩基配列を示す。

[図16-2]図16は、本発明に関連するアミノ酸配列又は塩基配列を示す。

[図16-3]図16は、本発明に関連するアミノ酸配列又は塩基配列を示す。

[図16-4]図16は、本発明に関連するアミノ酸配列又は塩基配列を示す。

[図16-5]図16は、本発明に関連するアミノ酸配列又は塩基配列を示す。

[図16-6]図16は、本発明に関連するアミノ酸配列又は塩基配列を示す。

[図16-7]図16は、本発明に関連するアミノ酸配列又は塩基配列を示す。

[図16-8]図16は、本発明に関連するアミノ酸配列又は塩基配列を示す。

[図16-9]図16は、本発明に関連するアミノ酸配列又は塩基配列を示す。

[図16-10]図16は、本発明に関連するアミノ酸配列又は塩基配列を示す。

[図17]図17は、Enko B蛋白質と細胞質雄性不稔（CMS）遺伝子を含むRNAとの結合解析を示す。

[図18]図18は、ORF687様蛋白質とRNAとの結合を示す。

[図19]図19は、オグラ型細胞質に働く稔性回復因子の結合配列予測を示す。

[図20]図20は、ORF687様蛋白質の候補結合RNA領域の二次構造と構造変化を示す。

[図21-1]図21は、ORF687様蛋白質のアライメントを示す。

[図21-2]図21は、ORF687様蛋白質のアライメントを示す。

[図22]図22は、様々なダイコン品種に含まれるORF687様蛋白質の塩基指定アミノ酸の一覧を示す。

発明を実施するための形態

[0019] [PPR モチーフ及びPPR蛋白質]

本発明で「PPRモチーフ」というときは、特に記載した場合を除き、Web上の蛋白質ドメイン検索プログラムでアミノ酸配列を解析した際に、PfamにおいてPF01535、PrositeにおいてPS51375で得られるE値が所定値以下（望ましくはE-03）のアミノ酸配列をもつ30～38アミノ酸で構成されるポリペプチドをいう。本発明で定義するPPRモチーフを構成するアミノ酸の位置番号は、PF01535とほぼ同義である一方で、PS51375のアミノ酸の場所から2引いた数（例；本発明の1番→PS51375の3番）に相当する。ただし、“ii”（-2）番のアミノ酸というときは、PPRモチーフを構成するアミノ酸の後ろ（C末端側）から2番目のアミノ酸、又は次のPPRモチーフの1番アミノ酸に対して2コN末端側、すなわち-2番目のアミノ酸とする（図1）。次のPPRモチーフが明確に同定されない場合、次のヘリックス構造の1番目のアミノ酸に対して、2コ前のアミノ酸を“ii”とする。Pfamについては<http://pfam.sanger.ac.uk/>、Prositeについては、<http://www.expasy.org/prosite/>を参照することができる。

[0020] PPRモチーフの保存アミノ酸配列は、アミノ酸レベルでの保存性は低い、2次構造上で2つの α ヘリックスはよく保存されている。典型的なPPRモチーフは35アミノ酸で構成されるが、その長さは30～38アミノ酸と可変的である。

[0021] 本発明でいうPPRモチーフは、より具体的には、式1で表される、30～38アミノ酸長のポリペプチドからなる。

[0022] [化4]

(Helix A)-X-(Helix B)-L (式1)

式中：

Helix Aは、12アミノ酸長の、 α ヘリックス構造を形成可能な部分であって、式2で表され、

[0023] [化5]

$A_1-A_2-A_3-A_4-A_5-A_6-A_7-A_8-A_9-A_{10}-A_{11}-A_{12}$ (式2)

式2中、 $A_1 \sim A_{12}$ はそれぞれ独立にアミノ酸を表し；

Xは、存在しないか又は1～9アミノ酸長からなる部分であり；

Helix Bは、11～13アミノ酸長からなる、 α ヘリックス構造を形成可能な部分であり；

Lは、2～7アミノ酸長の、式3で表される部分であり；

[0024] [化6]

$L_{vii}-L_{vi}-L_v-L_{iv}-L_{iii}-L_{ii}-L_i$ (式3)

式3中、各アミノ酸は、“i” (-1)、“ii” (-2)、とC末端側からナンバリングされ、

ただし、 $L_{iii} \sim L_{vii}$ は存在しない場合がある。

[0025] 本発明で「PPR蛋白質」というときは、特に記載した場合を除き、上述のPPRモチーフを、1個以上、好ましくは2個以上有するPPR蛋白質をいう。本明細書で「蛋白質」というときは、特に記載した場合を除き、ポリペプチド（複数のアミノ酸がペプチド結合した鎖）からなる物質全般をいい、比較的低下

子のポリペプチドからなるものも含まれる。本発明で「アミノ酸」という場合、通常のアミノ酸分子を指すことがあるほか、ペプチド鎖を構成しているアミノ酸残基を指すことがある。いずれを指しているかは、文脈から、当業者には明らかである。

[0026] PPR蛋白質は植物に多く存在し、シロイヌナズナでは500蛋白質、約5000モチーフが見いだせる。イネ、ポプラ、イワヒバ等、多くの陸上植物にも多様なアミノ酸配列のPPRモチーフ及びPPR蛋白質が存在する。いくつかのPPR蛋白質は、花粉形成（雄性配偶子）の形成に働く稔性回復因子として、雑種強勢のためのF1種子取得のための重要な遺伝子であることが知られている。稔性回復と類似して、いくつかのPPR蛋白質は種分化に作用していることが明らかになっている。ほとんどのPPR蛋白質は、ミトコンドリアか葉緑体中のRNAに働くことも分かっている。

[0027] 動物では、LRPPRCと同定されるPPR蛋白質の異常がLeigh syndrom French Canadian (LSFC; リー症候群、亜急性壊死性脳脊髄症)を引き起こすことが知られている。

[0028] 本発明で、PPRモチーフのRNA塩基との結合性に関し、「選択的」というときは、特に記載した場合を除き、RNA塩基のいずれか一つの塩基に対する結合活性が、他の塩基に対する結合活性より高いことをいう。この選択性は、当業者であれば実験を企画し、確認することができるほか、本明細書の実施例に開示されているように、計算により求めることもできる。

[0029] 本発明でRNA塩基というときは、特に記載した場合を除き、RNAを構成するリボヌクレオチドの塩基を指し、具体的には、アデニン(A)、グアニン(G)、シトシン(C)、又はウラシル(U)のいずれかをいう。なおPPR蛋白質は、RNA中の塩基に対して選択性を有しうるが、核酸モノマーに結合するわけではない。

[0030] 本発明以前には、PPRモチーフとしての保存アミノ酸の配列検索法は確立しているが、選択的なRNA塩基との結合に関する法則性は、まったく発見されていないかった。

[0031] 本発明により、以下の知見が提供される。

[0032] (I) 選択的結合のために重要なアミノ酸の位置に関する情報。具体的には、PPRモチーフの、1、4、“ii”(-1)番の3つのアミノ酸の組み合わせ (A_1 、 A_4 、 L_{ii})、又は4、“ii”(-1)番の2つのアミノ酸の組み合わせ (A_4 、 L_{ii}) が、RNA塩基との選択的な結合のために重要であり、これらの組み合わせにより、結合するRNA塩基がいずれであるかを決定できる。

[0033] 本発明は、本発明者らにより見出された、 A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせ、及び／又は A_4 、及び L_{ii} の2つのアミノ酸の組み合わせに関する知見に基づく。

[0034] (II) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせとNA塩基との対応に関する情報。具体的には、下記のようなものである。

(3-1) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、バリン、アスパラギン及びアスパラギン酸の場合、そのPPRモチーフは、Uに強く結合し、次にCに、その次にA又はGに対して結合するという、選択的なRNA塩基結合能を有する。

(3-2) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、バリン、トレオニン、アスパラギンの場合、そのPPRモチーフは、Aに強く結合し、次にGに、その次にCに対して結合するが、Uには結合しないという、選択的なRNA塩基結合能を有する。

(3-3) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、バリン、アスパラギン、アスパラギン酸の場合、そのPPRモチーフは、Cに強く結合し、次にA又はUに対して結合するが、Gには結合しないという、選択的なRNA塩基結合能を有する。

(3-4) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、グルタミン酸、グリシン、アスパラギン酸の場合、そのPPRモチーフは、Gに強く結合するが、A、U及びCには結合しないという、選択的なRNA塩基結合能を有する。

(3-5) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、イソロイシン、アスパラギン、アスパラギン酸の場合、そのPPRモチーフは、Cに強く結合し

、次にUに、その次にAに対して結合するが、Gには結合しないという、選択的なRNA塩基結合能を有する。

(3-6) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、バリン、トレオニン、アスパラギン酸の場合、そのPPRモチーフは、Gに強く結合し、次にUに対して結合するが、AとCには結合しないという、選択的なRNA塩基結合能を有する。

(3-7) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、リジン、トレオニン、アスパラギン酸、の場合、そのPPRモチーフは、Gに強く結合し、次にAに対して結合するが、U及びCには結合しないという、選択的なRNA塩基結合能を有する。

(3-8) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、フェニルアラニン、セリン、アスパラギンの場合、そのPPRモチーフは、Aに強く結合し、次にCに、その次にG及びUに対して結合するという、選択的なRNA塩基結合能を有する。

(3-9) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、バリン、アスパラギン、セリンの場合、そのPPRモチーフは、Cに強く結合し、次にUに対して結合するが、A及びGには結合しないという、選択的なRNA塩基結合能を有する。

(3-10) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、フェニルアラニン、トレオニン、アスパラギンの場合、そのPPRモチーフは、Aに強く結合するが、G、U及びCには結合しないという、選択的なRNA塩基結合能を有する。

(3-11) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、イソロイシン、アスパラギン、アスパラギン酸の場合、そのPPRモチーフは、Uに強く結合し、次にAに対して結合するが、G及びCには結合しないという、選択的なRNA塩基結合能を有する。

(3-12) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、トレオニン、トレオニン、アスパラギンの場合、そのPPRモチーフは、Aに強く結合する

が、G、U及びCには結合しないという、選択的なRNA塩基結合能を有する。

(3-13) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、イソロイシン、メチオニン、アスパラギン酸の場合、そのPPRモチーフは、Uに強く結合し、次にCに対して結合するが、A及びGには結合しないという、選択的なRNA塩基結合能を有する。

(3-14) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、フェニルアラニン、プロリン、アスパラギン酸の場合PPR、そのモチーフは、Uに強く結合し、次にCに対して結合するが、A及びGには結合しないという、選択的なRNA塩基結合能を有する。

(3-15) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、チロシン、プロリン、アスパラギン酸の場合、そのPPRモチーフは、Uに強く結合するが、A、G及びCには結合しないという、選択的なRNA塩基結合能を有する。

(3-16) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、ロイシン、トレオニン、アスパラギン酸の場合、そのPPRモチーフは、Gに強く結合するが、A、U及びCには結合しないという、選択的なRNA塩基結合能を有する。

[0035] (II) A_4 、及び L_{ii} の2つのアミノ酸の組み合わせとNA塩基との対応に関する情報。具体的には、下記のようなものである。

(2-1) A_4 、 L_{ii} が、順に、アスパラギン、アスパラギン酸の場合、そのPPRモチーフは、Uに強く結合し、次にC、その次にA及びGに対して結合するという、選択的なRNA塩基結合能を有する。

(2-2) A_4 、 L_{ii} が、順に、アスパラギン、アスパラギンの場合、そのPPRモチーフは、Cに強く結合し、次にU、その次にA及びGに対して結合するという、選択的なRNA塩基結合能を有する。

(2-3) A_4 、 L_{ii} が、順に、トレオニン、アスパラギンの場合、そのPPRモチーフは、Aに強く結合し、次にG、U及びCに対して弱く結合するという、選択的なRNA塩基結合能を有する。

(2-4) A_4 、 L_{ii} が、順に、トレオニン、アスパラギン酸の場合、そのPPRモチーフは、Gに強く結合し、次にA、U及びCに対して弱く結合するという、選択的

なRNA塩基結合能を有する。

(2-5) A_4 、 L_{ii} が、順に、セリン、アスパラギンの場合、そのPPRモチーフは、Aに強く結合し、次にG、U及びCに対して結合するという、選択的なRNA塩基結合能を有する。

(2-6) A_4 、 L_{ii} が、順に、グリシン、アスパラギン酸の場合、そのPPRモチーフは、Gに強く結合し、次にU、その次にAと結合するが、Cに結合しないという、選択的なRNA塩基結合能を有する。

(2-7) A_4 、 L_{ii} が、順に、アスパラギン、セリンの場合、そのPPRモチーフは、Cに強く結合し、次にU、その次にA及びGに対して結合するという、選択的なRNA塩基結合能を有する。

(2-8) A_4 、 L_{ii} が、順に、プロリン、アスパラギン酸の場合、そのPPRモチーフは、Uに強く結合し、次にG、C及びCに対して結合するが、Aに結合しないという、選択的なRNA塩基結合能を有する。

(2-9) A_4 、 L_{ii} が、順に、グリシン、アスパラギンの場合、そのPPRモチーフは、Aに強く結合し、次にGに対して結合するが、C及びUに結合しないという、選択的なRNA塩基結合能を有する。

(2-10) A_4 、 L_{ii} が、順に、メチオニン、アスパラギン酸の場合、そのPPRモチーフは、Uに強く結合し、次にA、G及びCに対して弱く結合するという、選択的なRNA塩基結合能を有する。

(2-11) A_4 、 L_{ii} が、順に、ロイシン、アスパラギン酸の場合、そのPPRモチーフは、Cに強く結合し、次にUに対して結合するが、A及びGに結合しないという、選択的なRNA塩基結合能を有する。

(2-12) A_4 、 L_{ii} が、順に、バリン、トレオニンの場合、そのPPRモチーフは、Uに強く結合し、次にAに対して結合するが、G及びCに結合しないという、選択的なRNA塩基結合能を有する。

[0036] 本明細書の実施例では、遺伝学的又は分子生物学的に、部分的には解析された蛋白質とその潜在的なRNA標的配列との結合を、さらに計算科学的に解析することにより、上記の知見を得ている。より詳細には、蛋白質とRNAとの結

合又はその選択的な結合を、P値（probability;確率）を指標に解析している。本発明では、一般的な有意水準であるP値が0.05以下(5%以下の偶然性)のとき、好ましくはP値が0.01以下（1%以下の偶然性）、のとき、より好ましくはそれより有意なP値が算出されたとき、その蛋白質とRNAとが結合する確率が十分に高い、と評価している。このようなP値による判断は当業者間であれば十分に理解できる。

[0037] 特定の位置のアミノ酸の特定の組み合わせとRNA塩基との結合性は、実験により確認することができる。このような目的での実験は、PPRモチーフ又は複数のPPRモチーフを含む蛋白質の調製、基質RNAの調製、及び結合性試験（例えば、ゲルシフト法）を含む。各々の実験は当業者にはよく知られているし、より具体的な手順・条件としては、例えば特許文献2を参考にすることができる。

[0038] [PPRモチーフ及びPPR蛋白質の利用]

同定及び設計：

一のPPRモチーフは、RNAの特定の塩基を認識しうる。そして、本発明に基づけば、特定の位置のアミノ酸を適切にすることで、A、U、G、Cそれぞれに選択的なPPRモチーフを選択又は設計することができ、さらにはそのようなPPRモチーフの適切な連続を含む蛋白質は、対応する特異的な配列を認識しうる。そのため、本発明に基づけば、特定の塩基配列を有するRNAに選択的に結合する天然型PPR蛋白質を予測・同定することができ、また逆に、PPR蛋白質の結合の標的となるRNAを予測・同定することができる。標的の予測・同定は、遺伝子的実体を明らかにするのに役立ち、また標的の利用可能性を拡大する点でも有用である。

[0039] さらに本発明により、所望のRNA塩基に選択的に結合可能なPPRモチーフ、及び所望のRNAに配列特異的に結合可能な、複数個のPPRモチーフを有する蛋白質を設計することができる。設計に際し、PPRモチーフ中の重要な位置のアミノ酸以外の部分は、天然型のPPRモチーフの配列情報を参考にすることができる。また、全体として天然型を用い、該当位置のアミノ酸だけを置換する

ことにより、設計してもよい。PPRモチーフの繰り返し数は、標的配列に応じ、適宜とすることができるが、例えば2個以上とすることができ、2～20個とすることができる。

[0040] 設計に際し、1、4、“ii”番アミノ酸の組み合わせ、又は4、“ii”番アミノ酸の組み合わせ以外に配慮してもよい。例えば、前掲特許文献2に記載された8番及び12番のアミノ酸の考慮は、RNA結合活性を呈するために重要である場合がある。本発明者らの検討に拠ると、あるPPRモチーフの A_8 と、それと同じPPRモチーフの A_{12} とが、RNA結合において協同している可能性がある。 A_8 は、塩基性アミノ酸、好ましくはリジン、又は酸性アミノ酸、好ましくはアスパラギン酸とすることができ、 A_{12} は、塩基性アミノ酸又は中性アミノ酸又は疎水性アミノ酸とすることができる。

[0041] 設計されたモチーフ又は蛋白質は、当業者にはよく知られた方法により調製することができる。すなわち、本発明は、1、4、“ii”番アミノ酸の組み合わせ、又は4、“ii”番アミノ酸の組み合わせに着目した特定のRNA塩基に選択的に結合するPPRモチーフ、及び特定配列を有するRNAに特異的に結合するPPR蛋白質を提供する。中でも、PPR蛋白質の機能として稔性に対する作用を考えた場合、上記3アミノ酸の組み合わせまたは2アミノ酸の組み合わせのいずれの場合であっても、4番アミノ酸(A_4)および“ii”番アミノ酸が有効であることが見出された。このようなモチーフ及び蛋白質は、当業者にはよく知られた方法で、比較的大量にも調製可能であり、そのような方法は、目的のモチーフ又は蛋白質が有するアミノ酸配列から、それをコードする核酸配列が決定し、クローニングし、目的のモチーフ又は蛋白質を生産する形質転換体を作製することを含む。

[0042] 複合体の調製及びその利用：

本発明により提供されるPPRモチーフ又はPPR蛋白質は、機能性領域を連結し、複合体とすることができる。機能性領域とは、生体内又は細胞内で特定の生物学的機能、例えば酵素機能、触媒機能、阻害機能、亢進機能などの機能を有する部分、又は標識としての機能を有する部分をいう。そのような領

域は、例えば、蛋白質、ペプチド、核酸、生理活性物質、薬剤からなる。機能性領域が蛋白質である場合の例は、リボヌクレアーゼ (RNase) である。RNaseの例は、RNase A (例えば、bovine pancreatic ribonuclease A: PDB 2AAS)、RNase Hである。このような複合体は天然には存在せず、新規なものである。

[0043] さらに本発明により得られる複合体は、生体内又は細胞内に、RNA配列特異的に機能性領域をデリバリーし、機能させることができる可能性がある。これにより、ジンクフィンガー蛋白質(前掲非特許文献1)やTAL effector(前掲非特許文献2、前掲特許文献1)と同様、生体内又は細胞内において、RNA配列特異的に改変・破壊を行うことができる可能性があり、また新たな機能を付与できる可能性がある。さらにはRNA配列特異的に薬物を送達し、することができる可能性がある。したがって、本発明は、RNA配列特異的な機能性物質の送達方法を提供するものでもある。

[0044] いくつかのPPR蛋白質は、花粉形成(雄性配偶子)の形成に働く稔性回復因子として、雑種強勢のためのF1種子取得において重要であることが知られている。本発明により、未だ同定されていない稔性回復因子を同定し、さらにその因子を高度に利用する技術の開発が期待できる。例えば、本願実施例において明らかになった事例において、細胞質雄性不稔性の稔性回復因子として働くPPR蛋白質遺伝子において、特定のPPRモチーフにおけるアミノ酸多型を検出することにより、当該遺伝子における多型と稔性との関連に基づいて、被検サンプルのPPR蛋白質遺伝子が稔性に関する遺伝型か、不稔に関連する遺伝型かを判定することができる。この場合、多型を検出する対象であるPPR蛋白質遺伝子としては、例えばダイコンの場合、園紅のORF687蛋白質(Enko Bと命名)をコードする「ORF687遺伝子」と相同な遺伝子座に座乗するファミリー遺伝子、Enko Bと90%以上のアミノ酸配列同一性を有する遺伝子、Enko Bをコードする「ORF687遺伝子」90%以上の塩基配列同一性を有する遺伝子、とが含まれる。ここで、園紅のORF687蛋白質(Enko Bと命名)をコードする「ORF687遺伝子」と相同な遺伝子座に座乗するファミリー遺伝子という場合

、図21及び図22に示す遺伝子（Kosena B、Comet B、Enko A、Comet A、Icicle CA、rrORF690-1、rrORF690-2、PC_PPR-A、PC_PPR-BL）がすべて含まれるが、これらには限定されない。また、Enko Bと90%以上のアミノ酸配列同一性を有する遺伝子、Enko Bをコードする「ORF687遺伝子」90%以上の塩基配列同一性を有する遺伝子は、遺伝子データベースを検索することにより、取得することができ、由来の種もダイコンには限定されない。またPPRモチーフとしては、上述した式1で表される30～38アミノ酸長のポリペプチドからなるPPRモチーフであり、PPR蛋白質は、その様なPPRモチーフを1個以上（好ましくは2～16個）含むことを特徴していてもよい。このPPRモチーフにおける多型という場合、それぞれのPPRモチーフでRNAへの結合を司っていることが本発明において明らかになった、1、4、“ii”番アミノ酸の組み合わせ、又は4、“ii”番アミノ酸の組み合わせに基づく多型を利用することができる。図4Bもしくは4Dで算出したP値が示す通り、このようなPPRモチーフのRNAへの結合を司っているアミノ酸の組み合わせのうち、4番アミノ酸がもっとも重要な役割を果たしており、次いで“ii”番アミノ酸が重要な役割を果たしている。そして、Enko BのPPR蛋白質と比較して、検定対象とする遺伝子のコードする蛋白質中のすべてのPPRモチーフの4番アミノ酸がEnko Bと同一であること、または対応するすべてのPPRモチーフの“ii”番アミノ酸がEnko Bと同一であることが、稔性回復因子としての機能に重要であることを明らかにした。また、稔性回復と類似して、いくつかのPPR蛋白質が種分化に作用していることが明らかになっている。当該PPR蛋白質の標的RNAの同定・改変は、今まで不可能だった種間の交配の可能性を期待させる。さらにほとんどのPPR蛋白質は、ミトコンドリア、葉緑体中のRNAに働くことから、本発明で提供される新規PPR蛋白質は、光合成、呼吸、有用代謝物の合成に関する機能の改変・向上に資するであろう。

[0045] 一方、動物では、LRPPRCと同定されるPPR蛋白質の異常がLeigh syndrom French Canadian (LSFC; リー症候群、亜急性壊死性脳脊髄症)を引き起こすことが知られている。本発明は、LSFCの処置（予防、治療、進行の抑制）に寄与

しうる。

[0046] また、PPR蛋白質は、オルガネラで見られるすべてのRNA加工のステップ、切断、RNA編集、翻訳、スプライシング、RNA安定化に関与している。本発明により、PPRモチーフの結合塩基選択性を改変することで、所望のRNAの発現を改変することが期待できる。

[0047] 本発明で材料にしたPPR蛋白質は、専らRNA編集（RNA上での遺伝情報の変換；多くの場合、C→U）の編集部位の指定に働く（後掲参照文献2及び3参照）。このタイプのPPR蛋白質は、RNA変種酵素と相互作用すると示唆される付加モチーフがC末端側に存在する。このような構造を有するPPR蛋白質により、塩基多型を導入すること、又は塩基多型に起因した疾患又は状態を処置することが期待できる。

[0048] また一部のPPR蛋白質のC末端側には、RNA切断酵素が付与されているものもある。N末端側のPPRモチーフの結合RNA塩基選択性を改変することで、RNA配列特異的なRNA切断酵素を構成しうる。また、GFP等の標識部分を連結した複合体は、所望のRNAを生体内で可視化するために用いうる。

[0049] 他方、既存のPPR蛋白質のなかには、DNAに作用するものがある。一つはミトコンドリア遺伝子の転写活性化因子であり、他の一つは核に局在する転写活性化因子であると報告されている。したがって、本発明で得られた知見を基に、所望のDNA配列に結合する蛋白質因子の設計も可能であろう。

実施例

[0050] 実施例1：RNA編集に関わるPPR蛋白質及びその標的配列の収集

図2に示した情報を参照し、これまでに解析されたシロイヌナズナのRNA編集に関わるPPR蛋白質（配列番号2～24）をシロイヌナズナゲノム情報データベース（MATDB: <http://mips.gsf.de/proj/thal/db/index.html>）、標的となるRNA編集部位の周辺配列（配列番号48、50、53、55、57、59、60、61、62、63、64、65、68、69、70、71、73、74、76、78、80、122、206、228、232、252、284、316、338、339、358、430、433、455、552、563）をRNA編集データベース（http://biologia.unical.it/py_script/overview.html）より収集し

た。RNA配列は編集されるC（シトシン）残基を含む、その上流31塩基を収集した。収集した全ての蛋白質と、それぞれの蛋白質に対応するRNA編集部位を図2に示した。

[0051] 蛋白質中のPPRモチーフ構造は、Uniprot データベース(<http://www.uniprot.org/>)の情報と共に、本発明で定義するアミノ酸番号を付与した。実験に用いたシロイヌナズナ24コ（配列番号2～25）のPPR蛋白質に含まれるPPRモチーフとそのアミノ酸番号を、図3に記す。

[0052] 実施例2：結合塩基選択性を付与するアミノ酸の同定

これまでの研究から、RNA編集に関わるPPR蛋白質はそのC末端側に特定の保存アミノ酸配列を持つモチーフ（E、E+及びDYWモチーフ、ただしDYWはしばしば存在しない）を持つことが分かっている。E+モチーフ中の十数アミノ酸は、RNAとの選択的な結合でなく、C（シトシン）からU（ウラシル）への変換に必要なことが示唆されている（参照文献3）。また、編集されるCの認識に必要な情報は、その上流20塩基及び下流5塩基に含まれていることが過去の非特許論文によって示唆されている。すなわち、PPR蛋白質中の複数個のPPRモチーフは、編集されるCの上流配列の「どこか」を認識し、E+モチーフが編集されるCの近傍に位置すると予想できる。さらに、PPRモチーフ中の特定のアミノ酸が結合する上流配列のRNA残基を認識する可能性が考えられる（図4A）。

[0053] この可能性を、実施例1で記したシロイヌナズナの24コのRNA編集PPR蛋白質とその標的RNA配列を用いて検証した。そこでまず、蛋白質中の最後のPPRモチーフを、編集されるCの一つ目の塩基に配置し、全てのPPRモチーフをRNA残基と、1対1の対応関係、かつ直線的な連続性、で整列させた（図4A、整列P1）。次にRNA配列を1塩基ずつ右側にずらすことで、P2～P6の整列を得た。この整列P1～P6のデータセットにおいて、それぞれのPPRモチーフと対応するRNA残基の情報を収集した。

[0054] 1箇所のRNA編集部位に働くPPR蛋白質に対しては、出現したRNA残基（A、U、G or C）に1点を付与した。2箇所、又は3箇所のRNA編集に働くPPR蛋白質に対しては、出現したRNA残基にそれぞれ0.5点、又は0.3点を付与した。次にPP

Rモチーフ中のアミノ酸番号ごとにアミノ酸の種類で並べ替えを行った。通常、アミノ酸の種類と出現するRNA塩基の種類は無作為 (high-randomness or high-entropy) であると予測できる (例、図4Aの右上図)。しかし、もし、特定の箇所のアミノ酸が結合RNA塩基選択能力を有するとすると、正しい整列 (上記P1~P6) において、対応するRNA塩基が一種類もしくは限定された種類に収束すると予測される (low randomness or low entropy; 例、図4Aの右下図)。

[0055] 上記で作成したP1~P6の整列のデータセットを対象に、PPRモチーフの全てのアミノ酸番号において、上記のlow randomnessを算出した。Low randomnessは、理論値 (全ての塩基出現頻度の平均値) に対するカイ二乗検定によって算出した (例、図5)。

[0056] その結果、P4の整列における1番、4番、及び“ii” (-2) 番アミノ酸において、有意値 $P < 0.01$ (1 %以下の確率 (probability)) が算出された (図4B)。すなわち、RNA編集PPR蛋白質中の最後のPPRモチーフが編集されるCの4コ前の塩基に配置され、3つのアミノ酸 (1番、4番、及び“ii” 番) が結合RNA塩基選択を司ることを示している。また、P3及びP5の整列において、有意なP値が算出されなかったことより、両脇のPPRモチーフからの干渉がない、すなわち、一つのPPRモチーフが一つのRNA塩基を認識し、モチーフ構成に依存しないことを示している。整列P4の他のアミノ酸、及び他の整列の全てのアミノ酸において、有意なP値は得られなかった (図6)。また、プリン (AとG) 又はピリミジン (CとU) (RY) でRNA塩基を分類し、同様の計算を行ったところ、4番アミノ酸のみで非常に有意なP値 ($P < 0.01$)、が得られた (図4C)。これは4番アミノ酸が結合するRNA塩基のプリン/ピリミジンを主に区別していることを示している。図4Cで示したPPRモチーフ中のRNA認識アミノ酸による結合塩基指定能力についてさらに詳細に解析を行った。その結果、4番アミノ酸が結合する塩基のプリン/ピリミジン (RY) を主に区別する以外に、“ii” (-2) 番アミノ酸が塩基のアミノ型 (AおよびC) ケト型 (GおよびU) (MK) の区別に働くことがわかった (図4D)。

[0057] 3回以上使用されている3つのアミノ酸（1番、4番、及び“ii”番）の組み合わせをPPRモチーフのRNA認識コードのうち、triPPR codeとして、それぞれのP値を算出し、その結合RNA塩基指定能力を算出した。同定したtriPPR codeの一部を図4Eに示した。

[0058] 当該3箇所のアミノ酸が非常に多様であるため、2箇所のアミノ酸（1&4、1&“ii”又は4&“ii”）での結合RNA塩基指定能力を算出したところ、4&“ii”番アミノ酸の組み合わせにおいて、顕著なP値が算出された（図7）。そのため、3回以上使用されている4&“ii”番アミノ酸の組み合わせを、PPRモチーフのRNA認識コードのうち、diPPR codeとした。同定したtriPPR code、diPPR codeを図8に示した。

[0059] 実施例3：同定したRNA認識コードの検証

シロイヌナズナのRNA編集PPR蛋白質を用いて同定したPPRモチーフのRNA認識コードの検証をおこなった。検証には、ヒメツリガネゴケのRNA編集PPR蛋白質を用いた。ヒメツリガネゴケ（以下、コケ）では、計13箇所（ミトコンドリア11箇所、葉緑体2箇所；配列番号32～44）のRNA編集が行われることが既に明らかになっている。さらに、6コのPPR蛋白質（PpPPR_56, 71, 77, 78, 79, and 91）が9箇所のRNA編集にそれぞれ働くことが明らかになっている。蛋白質と対応するRNA編集部位を図9に示した。

[0060] 検証は図10に示すとおりに行った。まず、コケPPR蛋白質のアミノ酸配列情報を非特許論文より取得し（配列番号26～31；図2及び図9）、図1に定義したPPRモチーフモデルに従って、3つのアミノ酸（1番、4番、及び“ii”番）をそれぞれのPPRモチーフから抽出した。抽出した3つのアミノ酸の組み合わせが、シロイヌナズナより同定したtriPPR codeと一致する場合、そのcodeが呈する結合塩基得点行列（socring matrix）に置換した。次に、triPPR codeで変換できなかったPPRモチーフのうち、diPPR codeと一致する場合、当該モチーフをdiPPR codeの結合塩基得点行列に置換した。平行して、RNA編集部位周辺配列を（編集されるCを3'末端とする31 merの配列）を非特許論文より取得し（配列番号32～44；図2、図9、及び図16）、図10に示すようなRNA配列の

数字行列に置換した。次に、上述の整列P4（最後のPPRモチーフが編集されるCの4つ前の塩基に対応）に従うように、蛋白質の結合塩基得点行列とRNA配列の数字行列のそれぞれのマス目通しでかけ算を行い、得られた値の和を蛋白質とRNA配列の適合値（Matching score）として算出した。この計算は、triPPR code、diPPR code、それぞれのPPR結合塩基得点行列（PPR scoring matrix）において行った。

[0061] 一つの蛋白質に対して、この計算を全てのコケのRNA編集部位（13箇所）に対して行った。また、RNA編集部位周辺配列の参照配列としてシロイヌナズナ葉緑体のRNA編集部位のRNA配列34箇所（図16、配列番号45～78）についても同様の計算を行った。

[0062] 次に、それぞれのRNA配列に対する蛋白質の適合値より、正規分布曲線を描き、それぞれのRNA配列に対する適合値の暫定P値をtriPPR code、及びdiPPR codeそれぞれ算出した。

[0063] 最終的なP値（蛋白質とRNA配列の適合値）は、triPPR codeとdiPPR codeの暫定P値の積として求めた。

[0064] それぞれのコケPPR蛋白質と13箇所のコケRNA編集部位との適合値を図11に示す。解析の結果、7種の蛋白質のうち、6種の蛋白質が正しいRNA編集部位に計算的に特定された。すなわち、この解析は、3つのアミノ酸（1番、4番、及び“ii”番）にPPRモチーフの結合RNA塩基指定に関する情報が全て含まれることを意味している。言い換えると、図8に示した3つ、又は2つ、の組み合わせのアミノ酸の情報（triPPR、diPPR code）を参照することで、意図するRNA配列に結合するPPR蛋白質を探索できることを示している。同時に、当該アミノ酸情報を有したPPRモチーフを用いること、又は連結することで、意図したRNA配列に結合する人工タンパク質を合成できることを示している。

[0065] 実施例4：未解析RNA編集PPR蛋白質の標的分子の同定

次に、コケより多くのRNA編集部位を含むシロイヌナズナを用いた解析を行った（葉緑体ゲノム34箇所（配列番号45～78）、ミトコンドリアゲノム488箇所（配列番号79～566）、図6を参照）。予測精度を検証するために、コード

抽出に用いた24種のPPR蛋白質のRNA変種部位予測を行った。その結果、葉緑体局在PPR蛋白質では、13個中10個が最低1個の正しいRNA編集部位を最も高いP値で予測した。ミトコンドリア局在PPR蛋白質では、11個中8個が正しいRNA編集部位をトップ20以内で予測した（図12）。この予測精度検証を基に、機能未知PPR蛋白質の標的RNA編集部位の予測を行った。AHG11変異体は、アブシジン酸経路に異常をきたす変異体であり、その遺伝子（ahg11、at2g44880）がコードする蛋白質は典型的なRNA編集PPR蛋白質様のモチーフ構造を有する（図13；配列番号1）。RNA編集部位を予測し、トップ20を含むミトコンドリア405箇所、葉緑体30箇所のRNA編集を実験的に検証した。その結果、7番目に高いP値で予測されたミトコンドリアnad4_376のRNA編集のみが変異体で異常をきたしていることが明らかになった（図13）。

[0066] 次に、オルガネラ全ゲノム配列、すなわち約 3×10^5 のRNA配列のデータセット、からの標的RNA配列の同定を試みた。この解析には、図8に示すPPR codeの蓋然性行列（probability matrix）を用いた。また、diPPR、triPPRコードと一致しないアミノ酸の組み合わせを持つモチーフには、background frequencyを適応した。作製した蛋白質の蓋然性行列は、シロイヌナズナの葉緑体全塩基配列（AP000423）と共に、MEME suiteのFIMO解析（http://meme.nbcr.net/meme4_6_1/fimo-intro.html）に供した。

[0067] その結果、CRR4とCRR21において、その標的RNA配列を正確に予測することが出来た。また、コケPPR蛋白質からもPPRコードの抽出を行うことでコードの改良を行ったところ（図15）、いくつかの蛋白質に置いて、その予測精度が大きく向上した。

[0068] これらの結果から、同定したPPRコードを用いることで、数十万パターンのRNA配列から一カ所の正しい標的配列を同定できることを示している。逆に言えば、コードに沿ったアミノ酸を当該位置（1、4、“ii”）に持つPPRモチーフ、を探索することで、意図する有用なRNA配列に結合する蛋白質を同定できる。もしくはPPRモチーフを連結させることで、高い配列選択性を有する人工RNA結合蛋白質を創出できることを示している。変異導入により、当該位置の

アミノ酸をPPRコードに準じた組み合わせにすることで、意図したRNA結合選択性を獲得させることが可能なことも当業者であれば理解できる。

[0069] 図15では、triPPR code、diPPR codeそれぞれの結合RNA塩基選択能力をP値で評価した。有意なP値 ($P < 0.05$) を示したPPRcodeは、高い結合RNA塩基選択能力を有すると推認できる。

[0070] 実施例5：ダイコンRfの標的RNA配列の予測

次に本発明で得られた知見を基に、細胞質雄性不稔性の稔性回復因子として働くPPR蛋白質の機能判定を行った（実施例5～9）。

[0071] 細胞質雄性不稔（Cytoplasmic Male Sterility；CMS）は、細胞質のゲノム、特にミトコンドリアのゲノムの変異が原因で雄性配偶子が正常に機能しなくなる形質である。この形質は、しばしば核に存在する稔性回復遺伝子（Restorer of Fertility；Rf）によって打ち消され、雄性配偶子が正常になることが知られている。一代雑種育種法に利用されており、農業上重要な形質の一つである。このCMS-Rfシステムにおいて、Rf遺伝子は多くの場合、PPR蛋白質をコードすることが知られている。

[0072] ダイコンやナタネの一代雑種育種法に用いられるオグラ型（別名、コセナ型）細胞質は、ミトコンドリアゲノムのorf125遺伝子の発現に由来し、核コードのorf687遺伝子の存在によって不稔性が解除され、可稔となる。orf687遺伝子産物はPPR蛋白質であり、orf125を含むRNAに作用することで、その発現を不活化し、結果として不稔性が解除されと考えられている。

[0073] しかし、様々なダイコン系統が有するorf687様遺伝子にはアミノ酸多型が有ること、このアミノ酸多型が稔性回復因子としての遺伝子の機能性に影響を及ぼすことがこれまでの育種学的解析から明らかになってきた。しかし、当該遺伝子のアミノ酸配列からその機能性を類推する手法は確立されていない。

[0074] そこで、まず優性Rfとして働くことが知られているダイコン品種、園紅のORF687蛋白質（Enko Bと命名）のアミノ酸から、PPRモチーフを特定し、塩基指定能力を司るアミノ酸（1、4、ii）を抽出し、PPRコードに変換後、ミトコ

ンドリアorf125を含む転写物に対して、標的RNA配列予測を行った（図19）。

[0075] 平行して、優性Rfとして働くことが知られているダイコン品種、園紅のORF687蛋白質（Enko Bと命名）、同じく園紅中に含まれ、ORF687と良く似ているが、劣性の遺伝子として働くORF687様蛋白質（enko Aと命名）、異なるダイコン品種である小瀬菜ゲノム中に存在する園紅ORF687と相同な遺伝子（kosen a Bと命名；劣性の遺伝子）、の3種のORF687様蛋白質を材料にその特徴を生化学的に解析した。

[0076] (5-1) ダイコンからのゲノムDNAの調製

ダイコンをムラシゲ・スクーク培地（2%ショ糖、0.5% Gellangamを含む）で3週間培養した。培養した植物の緑葉（0.5 g）をフェノール/クロロホルム抽出した後、エタノールを加えてDNAを不溶化した。回収したDNAを100 μ lのTE液（10 mM トリス・塩酸（pH 8.0）、1 mM EDTA）に溶解し、10ユニットのRNase A（DNase-free、タカラバイオ社）を加えて、37°Cで30分反応させた。その後、反応液を再度フェノール/クロロホルム抽出した後、エタノール沈殿によりDNAを回収した。10 μ gのDNAが得られた。

[0077] (5-2) ORF687様蛋白質をコードする遺伝子のクローニング

ダイコンゲノムDNAを鋳型に、Enko Bはオリゴヌクレオチドプライマー（Enko_B-FプライマーとEnko_B-Rプライマー；それぞれ配列番号567、568に記載）、Kosena Bはオリゴヌクレオチドプライマー（kosena_B-Fプライマーとkosena_B-Rプライマー；それぞれ配列番号569、570に記載）、Enko Aはオリゴヌクレオチドプライマー（Enko_A-FプライマーとEnko_A-Rプライマー；それぞれ配列番号571、572に記載）、を用いて、50 μ lの反応液を95°C 30秒、60°C 30秒、72°C 30秒の25サイクルでKOD-FX（TOYOBO社）をDNA伸長酵素として用い、PCRすることによって、それぞれ増幅した。

[0078] 得られたDNA断片は、pBAD/Thio-TOPO ベクター（Invitrogen社）を用いて、製品に添付するプロトコールに従ってクローニングした。DNA配列を決定し、目的と相当するDNA配列と相同な配列（Enko B、配列番号：573；kosen a B、配列番号：574；Enko A、配列番号：575）であることを確認した。

[0079] (5-3) 組換えORF687様蛋白質の調製

上で得られたプラスミドをEscherichia coli TOP10株 (Invitrogen社) に形質転換した。この大腸菌をアンピシリンが100 µg/mlの濃度で存在するLB培地300 ml (300 ml培地を含む1 L三角フラスコ) 中で、37℃で培養した。培養液の濁度が波長600 nmでの吸光度が0.5に達した時に、誘導物質であるL-アラビノースを最終濃度が0.2%になるように添加し、さらに4時間培養を行った。

[0080] 遠心による集菌後、菌体を1 mg/mlのリゾチームを含む200 mlのバッファー A (50 mM トリス・塩酸 pH 8.0、500 mM KCl、2 mM イミダゾール、10 mM MgCl₂、0.5% Triton X100、10% グリセロール) に懸濁し、超音波破碎と凍結溶解により菌体を破壊した。15,000×g、20分間の遠心分離後に、上清を粗抽出液として回収した。

[0081] この粗抽出液をバッファーAで平衡化したニッケルカラム樹脂 (ProBond A、Invitrogen社) を充填したカラムに供した。

[0082] カラムクロマトグラフィーは、20 mM イミダゾールを含むバッファーAで十分に洗浄した後、200 mM イミダゾールを含むバッファーAで目的タンパク質を溶出する二段階濃度勾配により行った。得られたタンパク質は、配列番号 (Enko B、配列番号 : 576 ; kosena B、配列番号 : 577 ; Enko A、配列番号 : 578) に記載のアミノ酸配列を備えるとともに、N末端側に溶解性を高めるためのチオレドキンのアミノ酸配列、C末端側にヒスチジンタグ配列を備える融合タンパク質である。精製画分100 µlを500 mLのバッファー E (20 mM トリス・塩酸 pH 7.9、60 mM KCl、12.5 mM MgCl₂、0.1 mM EDTA、17% グリセロール、2 mM DTT) で透析した後、精製標品とした。

[0083] (5-4) 基質RNAの調製

基質RNAとして、オグラ型細胞質ダイコンのミトコンドリアDNAの配列をふくむ3種のRNA、RNAa、RNAb、およびRNAcを用いた。

[0084] RNAaはオリゴヌクレオチドプライマーA-FプライマーとA-Rプライマー (それぞれ、配列番号579、580) 、RNAbはオリゴヌクレオチドプライマーB-Fプラ

イマーとB-Rプライマー（それぞれ、配列番号581、582）、RNACはオリゴヌクレオチドプライマーC-FプライマーとC-Rプライマー（配列番号583、584）、を用いて、上記のオグラ型細胞質ダイコンDNA 10 ngを鋳型DNAとして含む50 μ lの反応液を95°C30秒、60°C30秒、72°C30秒の25サイクルでKOD FX（TOYOBO社）をDNA伸長酵素として用い、PCRすることによって増幅した。それぞれのフォワードプライマー（-F）には、基質RNAを試験管内で合成するためのT7プロモータ配列を付加した。

[0085] 得られたDNA断片は、アガロースゲルで展開後、ゲルから切り出すことによって精製した。精製DNA断片を鋳型にNTP mix（10 nmol GTP、CTP、ATP、0.5 nmol UTP）、4 μ l [^{32} P] α -UTP（GEヘルスケア社、3000 Ci/mmol）、T7 RNA polymerase（タカラバイオ社）を含む20 μ lの反応液を37°C60分間反応させることで、基質RNAを合成した。

[0086] 基質RNAはフェノール/クロロホルム抽出、エタノール沈殿後、全量を6 M尿素を含む変性6%ポリアクリルアミドゲル電気泳動で展開し、X線フィルムで60秒間感光させることによって、 ^{32} P標識RNAを検出した。

[0087] 次に、 ^{32} P標識RNAをゲルから切り出し、200 μ lのゲル溶出液（0.3 M酢酸ナトリウム、2.5 mM EDTA、0.01%SDS）中に、4°Cで12時間浸し、RNAをゲルから溶出した。RNAのうち、1 μ lの放射活性を測定し、合成したRNAの総量を算出した。エタノール沈殿後、2500 cpm/ μ l（1 fmol/ μ l）になるように、RNAを超純水に溶解した。この調製方法で通常、2500 cpm/ μ lのRNAが約100 μ l得られた。

[0088] （5-5）蛋白質とRNAとの結合実験

Enko B（Rf）、Kosena B（rf）、およびEnko A（rf；園紅品種に存在する0 RF687様蛋白質）の組換え蛋白質を作製し、そのRNA結合活性を検証した。

[0089] 作製した組換え蛋白質（Enko B（配列番号：576）、Kosena B（配列番号：577）、Enko A（配列番号：578））のRNA結合活性は、ゲルシフト法によって解析した。反応液（10 mMトリス・塩酸 pH 7.9、30 mM KCl、6 mM MgCl₂、2 mM DTT、8%グリセロール、0.0067%Triton X-100）20 μ l中に上記の375 pM（

7.5 fmol/20 μ L) の基質RNA (BD120) と0~2500 nMの組換え蛋白質を混合し、25°Cで15分間反応した。その後、反応液に4 μ Lの80%グリセロール液を添加し、10 μ Lを1×TBE (89 mM Tris-HCl、89 mM Boric acid、2 mM EDTA) を含む10%未変性ポリアクリルアミドゲルで展開し、電気泳動後にゲルを乾燥させた。

[0090] ゲル中のRNAの放射活性をバイオイメージングアナライザーBAS2000 (フジフィルム社) で測定した。

[0091] 実施例6：組換え蛋白質を用いたRNA結合実験

図17は、Enko B蛋白質と細胞質雄性不稔 (CMS) 遺伝子を含むRNAとの結合解析を示す図である。このうち、図17Aは、ミトコンドリアorf125 近傍の模式図を示し、あわせて結合実験に用いたRNA a、RNA bc、RNA b、およびRNA cの領域を模式図で示した。図17Bは、Enko B蛋白質のRNA結合について示す図である。Enko Bタンパク質 (1.4 nmol) と³²P標識したRNA bc (0.1 ng) と共に、未標識のRNA a、RNA bc、RNA b、RNA c (RNA bc に対して、×5、×10 w/w；競合阻害物質として使用) を20 μ Lの反応液中で反応させ、ゲルシフト競合実験を行った。図左のComplex (▽) はタンパク質とRNAとの複合体を、Free (▼) はRNA のみを示す。

[0092] これらの図に示すように、蛋白質とRNAの結合は、³²P標識RNAの移動度の違いとして現れる。³²P標識RNA・蛋白質複合体の分子量が、³²P標識RNA単体の分子量より大きいため、電気泳動での移動度が遅くなるためである。

この実験では、EnkoBの組換え蛋白質を調製し、orf125を含むミトコンドリアRNAとの結合を競合ゲルシフト法で検証した。RI標識したRNAbと蛋白質を混合し、次に未標識RNAを添加した。すなわち、Complexで示した位置のバンドのシグナル強度がより減少したほうがcompetitorとして加えたRNAと蛋白質が結合する、すなわち、EnkoBが高い親和性で結合するRNA領域、であることを意味する。その結果、RNAbの領域にEnkoBは強く結合することが明らかとなった。

[0093] No. 208の候補配列は図19に示す結合配列予測で一番有意なP値を示し、tRNA

メチオニンの3'末端に正確に位置する。しかし、これまでの解析で、tRNA量およびorf125を含むRNAの形状（切断の有無）が不稔および回復系統において差が無いこと、in vitroの結合実験（図17B）でNo.208を含むRNAa配列とEnko Bが結合しないことから、この領域は、オグラ型細胞質の稔性・不稔性には関係ないと判断した。

[0094] よって、RNAb中に含まれるNo.316、352、373の領域に着目して解析を進めた。RNAbは125bから成る。scanning mutationにより、20b単位まで結合領域を絞り込もうと試みたが、単一の箇所には絞り込むことができなかった（データ未公開）。そのため、Enko Bの結合箇所は、RNAb中に複数ある可能性が考えられた。

[0095] 実施例7：Rf様蛋白質のRNA結合活性

図18は、ORF687様蛋白質とRNAとの結合を示す。このうち、図18Aは、ORF687様蛋白質のRNA結合特性に関して、Enko B (Rf)、Kosena B (rf)、Enko A (rf) とRNAbとの結合をゲルシフト法で解析した結果を示す。図18Bは、(A)の結果をグラフ化したものであり、このグラフより、各蛋白質のRNA結合能力を表す解離定数 (KD) を算出した。図18Cは、Enko B (Rf)、Kosena B (rf)、Enko A (rf) と潜在的な結合領域との適合度を図19と同様な方法で算出した。

[0096] その結果、非競合状態では、3つの蛋白質 (Enko B、Kosena B、Enko A) ともに、RNAbに高い親和性でRNAと結合した。Kosena Bに関しては競合状態でのRNA結合活性を解析したが、Enko Bと明確な差は見られなかった（図18Aおよび図18B）。

[0097] Kosena BはしばしばEnko Bより若干低いRNA結合活性を示す (KDで約2倍)。しかし、一般的なRNA結合における活性の強弱は10倍以上の差で検出されることが多く、この差が有意とは見なせない。

[0098] PPRコードに基づいた予測でも該当領域に対する適合スコアは、各蛋白質間で明確な差はない（図18C）。このことから、EnkoBとkosenaBの差は、単純なRNA結合親和性の違いではなく、結合した後の作用に違いがある可能性を検討

することとした。

[0099] さらに、オグラ型細胞質に働く稔性回復因子の結合配列予測を図19に示す。ここで、図19Aは、PPRコードを用いたEnko B蛋白質の結合予測を示しており、図19A下図にCMS遺伝子orf125を含むRNAの構造を示す。図19A中のRNAa～RNAcの領域は図17を参照。この図19Aでは、有意な高いP値を示した領域のうち、No.208、230、316、352、373に着目した（図19A）。

[0100] 次に、図19BにORF687蛋白質配列から予測された標的RNA配列（有意なP値を示した領域（No.208、316、352、373）の配列）のロゴ表記、候補となる結合RNA配列、劣性rfを持つダイコン品種、小瀬菜のORF687様蛋白質（Kosena B）配列から予測された標的RNA配列のロゴ表記、を示した。また、劣性rfであるKosena Bの予測結合塩基も示した。

[0101] EnkoBとKosenaBは、2及び3番目のPPRモチーフ中のアミノ酸多型により、指定塩基が異なることが明らかとなった（RfはUA、rfはGC）。この違いが、Rfとrfの機能的差異に直結すると予測できた。

[0102] 実施例8：RNAの構造予測と解析

コンピュータ予測およびin vitro RNA結合実験より、RfはRNAbの領域、特にNo.316、352、373、に結合する可能性が考えられた。in vitroの解析からRNAb中に複数箇所の結合個所がある可能性も考えられた。そこで、RNAb配列の2次構造予測を行い、該当領域に着目した。

[0103] 結果を図20に示す。ここで、図20は、ORF687様蛋白質の候補結合RNA領域の二次構造と構造変化を示す。図20Aは、No.306を含む領域の2次構造とORF687様蛋白質の予測結合サイトを示しており、各PPRモチーフをボックスで対応する塩基と共に示した。EnkoB(Rf)とKosena B (rf) で顕著な差がある2番目、3番目のPPRモチーフを強調した。図20Bは、No.352および373を含む領域の2次構造とORF687様蛋白質の予測結合サイトを示した。図20Cは、RNAbのEnko Bによる構造変化しており、RNAbとEnko B蛋白質を混合し、その後、2本鎖選択的なRNA分解酵素（Rnase V1）を添加した。

[0104] その結果、No.316領域は、orf125の開始コドン直下のステムループ構造に

相当することが明らかとなった（図20A）。また、Enko BとKosena Bとで多型が見られる2番および3番PPRモチーフはステムループの根元の二本鎖に位置した。特に3番PPRモチーフの対応塩基はEnko BではAで有るのに対して、Kosena Bでは、Cとなっている（図19Bを参照）。これらのことから、EnkoBが当該領域に結合し、ステムループ構造形成を促進することで、orf125の翻訳を阻害する作業仮説が考えられた。

[0105] No. 352および373領域でも2本鎖構造が予測され、Rf蛋白質が両側に結合することが考えられた（図20B）。しかし、この場合、Rf結合により、構造の破壊（1本鎖形成促進）が予測される。また、Rfとrfの違いである2番および3番PPRモチーフに対応する塩基と構造の違いは考えられず、具体的な分子機構の予測は出来なかった。

[0106] そこで、内部ラベルRNAを蛋白質と混合し、そこに、RNaseV1を添加し、標識RNAを限定分解した。RNaseV1はRNAの2本鎖領域のみを選択的に切断するRNAseである。その結果、蛋白質存在下で、基質RNAが早く分解されること、すなわち、Rf（Enko B）存在下で、2本鎖RNA形成が促進することが示された（図20C）。すなわち、Rfによるorf125 mRNAの2本鎖RNA形成による翻訳阻害が、オグラ型細胞質雄性不稔性の稔性回復の主原因と考えられた。

[0107] 実施例9：ORF687様遺伝子の稔性回復能についての機能判定

これまでに様々なダイコン品種よりORF687様遺伝子が単離されており、交配実験により、そのRfとしての機能性が類推されている。しかし、それぞれのアミノ酸配列は非常に似ており、全体のアミノ酸保存性からRfとしての機能性を判定することは出来ない。

[0108] 本実施例においてはまず、ORF687様蛋白質の配列解析を行った。具体的には、配列番号576～578、および585～591に示す蛋白質配列を材料にPPR蛋白質としての配列解析を行った。全ての配列を問い合わせ配列に用いて、CLUSTAL W (<http://www.genome.jp/tools/clustalw/>) によって、配列アライメントを得た。Web上のドメイン解析ソフト、

Pfam (<http://pfam.sanger.ac.uk/>)、

InterProScan (<http://www.ebi.ac.uk/Tools/InterProScan/>)、
Prosit (<http://www.expasy.org/prosit/>)、
を用いて、ORF687様蛋白質のアライメントを作成し、そしてそれぞれの蛋白質のPPRモチーフ構造を解析した。その結果を図21に示す。全てのORF687様遺伝子は16個のPPRモチーフから構成されている（図21）。

[0109] 得られたPPRモチーフモデルと非特許文献5に示されるアミノ酸番号に従って、1、2、” ii” (-2)番アミノ酸を抽出し、ORF様蛋白質の稔性回復能機能判定に用いた。

[0110] そこで、9種のRf様遺伝子の機能判定をPPRコードを用いて行った。前述のEnkoBと同じように塩基指定能力を司るアミノ酸（1、4、ii）を抽出し、PPRコードに変換後、アミノ酸種をRNA binding windowとして、その機能性判定に用いた（図22）。Enko BとKosena Bは、全体で99.4%の相同性を示すが、RNA binding windowには2箇所のアミノ酸多型があり、これがORF687様遺伝子の稔性回復に対する優性/劣性に深く関与すると考えられた（非特許論文4）。一方、品種Comet中で、Enko Bと相同な遺伝子座に座乗する遺伝子Comet BはEnko Bと98.0%の相同性を示し、RNA binding windowは全く同一である。過去の交配試験で得られたCometBが優性遺伝子であるという知見を裏付けることができた。また、Enko AはEnko Bの近傍に座乗する重複遺伝子だが、RNA認識の観点からも劣性の遺伝子であることが示唆された。これらのデータから、ORF687様遺伝子の稔性回復に対する優性/劣性は、Enko BのRNA binding windowと比較した場合に、ORF687様遺伝子における対応するPPRモチーフすべてにおいて、塩基指定能力を司るアミノ酸（1、4、ii）が同一であること、特に同一の4番アミノ酸（A₄）を有すること、あるいは同一の“ii”番アミノ酸を有すること、が重要であることが示唆された。中でも、特に同一の4番アミノ酸（A₄）を有することが重要と考えられた。この点から、稔性に関する情報が未知である様々な系統のダイコンにおいて、Enko Bと相同な遺伝子座に座乗する遺伝子、rrORF690-1、rrORF690-2、icicle_pprCA、PC_PPR-A、PC_PPR-BLについては、優性遺伝子であるEnko Bとは異なったRNA binding windowを持ち

、これら遺伝子も劣性rfであると考えられた。

[0111] 以上の結果は、本発明で示すPPRコードは、稔性回復因子として働くような産業上有用なPPR蛋白質の機能判定を高速化することができる。当該技術はCMS-Rfシステムを利用した一代雑種育種法に新しい系統を適用するさいに、候補Rf遺伝子の配列から、その稔性回復能の有無を判定することが可能である。発明者らは21種の新規ダイコン品種のORF687様遺伝子の機能判定を行い、19種のORF様遺伝子の稔性回復能の優性/劣性の判定に成功している（データ未公開）。本技術はオグラ型細胞質のダイコンに限らず、PPR蛋白質をRfとする様々な細胞質、植物種に適用可能である。

[0112] [実施例で引用した論文]

参考文献1: Small, I.D., and Peeters, N. (2000). The PPR motif - a TPR-related motif prevalent in plant organellar proteins. *Trends Biochem. Sci.* 25, 46-47.

参考文献2: Lurin, C., Andres, C., Aubourg, S., Bellaoui, M., Bitton, F., Bruyere, C., Caboche, M., Debast, C., Gualberto, J., Hoffmann, B., et al. (2004). Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell* 16, 2089-2103.

参考文献3: Okuda, K., Myouga, F., Motohashi, R., Shinozaki, K., and Shikanai, T. (2007). Conserved domain structure of pentatricopeptide repeat proteins involved in chloroplast RNA editing. *Proc Natl Acad Sci USA* 104, 8178-8183.

参考文献4: Koizuka N, Imai R, Fujimoto H, Hayakawa T, Kimura Y, et al. (2003) Genetic characterization of a pentatricopeptide repeat protein gene, orf687, that restores fertility in the cytoplasmic male-sterile Kosen radish. *Plant J* 34: 407-415.

参考文献5: Nakamura T, Yagi Y, Kobayashi K (2012) Mechanistic insight into pentatricopeptide repeat proteins as sequence-specific RNA-binding proteins.

nding proteins for organellar RNAs in plants. Plant & Cell Physiology
53: 1171-1179

請求の範囲

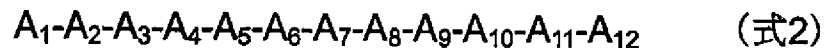
[請求項1] RNA塩基選択的に、又はRNA塩基配列特異的に結合可能な蛋白質を設計する方法であって；

蛋白質が、式1で表される30～38アミノ酸長のポリペプチドからなるPPRモチーフを1個以上（好ましくは2～14個）含む、蛋白質であり
[化1]



（式中：

Helix Aは、12アミノ酸長の、 α ヘリックス構造を形成可能な部分であって、式2で表され、
[化2]



式2中、 $\text{A}_1 \sim \text{A}_{12}$ はそれぞれ独立にアミノ酸を表し；
Xは、存在しないか又は1～9アミノ酸長からなる部分であり；
Helix Bは、11～13アミノ酸長からなる、 α ヘリックス構造を形成可能な部分であり；
Lは、2～7アミノ酸長の、式3で表される部分であり；
[化3]



式3中、各アミノ酸は、“i” (-1)、“ii” (-2)、とC末端側からナンバリングされ、

ただし、 $\text{L}_{\text{iii}} \sim \text{L}_{\text{vii}}$ は存在しない場合がある。）

A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせ、又は A_4 、 L_{ii} の2つのアミノ酸の組み合わせを、対象RNA塩基又は塩基配列に応じたものと

する、方法。

[請求項2]

A_1 、 A_4 、及び L_{ij} の3つのアミノ酸の組み合わせを、RNA塩基又は対象塩基配列に応じたものとする、請求項1に記載の方法であって、アミノ酸の組み合わせが、下記のいずれかに基づいて決定される、方法：

(3-1) A_1 、 A_4 、及び L_{ij} の3つのアミノ酸が、順に、バリン、アスパラギン、アスパラギン酸であるとき、そのPPRモチーフは、U（ウラシル）に選択的に結合でき；

(3-2) A_1 、 A_4 、及び L_{ij} の3つのアミノ酸が、順に、バリン、トレオニン、アスパラギンであるとき、そのPPRモチーフは、A（アデニン）に選択的に結合でき；

(3-3) A_1 、 A_4 、及び L_{ij} の3つのアミノ酸が、順に、バリン、アスパラギン、アスパラギンであるとき、そのPPRモチーフは、C（シトシン）に選択的に結合でき；

(3-4) A_1 、 A_4 、及び L_{ij} の3つのアミノ酸が、順に、グルタミン酸、グリシン、アスパラギン酸であるとき、そのPPRモチーフは、G（グアニン）に選択的に結合でき；

(3-5) A_1 、 A_4 、及び L_{ij} の3つのアミノ酸が、順に、イソロイシン、アスパラギン、アスパラギンであるとき、そのPPRモチーフは、C又はUに選択的に結合でき；

(3-6) A_1 、 A_4 、及び L_{ij} の3つのアミノ酸が、順に、バリン、トレオニン、アスパラギン酸であるとき、そのPPRモチーフは、Gに選択的に結合でき；

(3-7) A_1 、 A_4 、及び L_{ij} の3つのアミノ酸が、順に、リジン、トレオニン、アスパラギン酸であるとき、そのPPRモチーフは、Gに選択的に結合でき；

(3-8) A_1 、 A_4 、及び L_{ij} の3つのアミノ酸が、順に、フェニルアラニン、セリン、アスパラギンであるとき、そのPPRモチーフは、Aに選択的に結合でき；

(3-9) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸が、順に、バリン、アスパラギン、セリン、の場合であるとき、そのPPRモチーフは、Cに選択的に結合でき；

(3-10) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸が、順に、フェニルアラニン、トレオニン、アスパラギンであるとき、そのPPRモチーフは、Aに選択的に結合でき；

(3-11) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸が、順に、イソロイシン、アスパラギン、アスパラギン酸であるとき、そのPPRモチーフは、UまたはAに選択的に結合でき；

(3-12) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸が、順に、トレオニン、トレオニン、アスパラギンであるとき、そのPPRモチーフは、Aに選択的に結合でき；

(3-13) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸が、順に、イソロイシン、メチオニン、アスパラギン酸であるとき、そのPPRモチーフは、UまたはCに選択的に結合でき；

(3-14) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸が、順に、フェニルアラニン、プロリン、アスパラギン酸であるとき、そのPPRモチーフは、Uに選択的に結合でき；

(3-15) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸が、順に、チロシン、プロリン、アスパラギン酸であるとき、そのPPRモチーフは、Uに選択的に結合でき；

(3-16) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸が、順に、ロイシン、トレオニン、アスパラギン酸であるとき、そのPPRモチーフは、Gに選択的に結合できる。

[請求項3]

A_4 、及び L_{ii} の2つのアミノ酸の組み合わせを、RNA塩基又は対象塩基配列に応じたものとする、請求項1に記載の方法であって、アミノ酸の組み合わせが、下記のいずれかに基づいて決定される、方法：

(2-1) A_4 、 L_{ii} が、順に、アスパラギン、アスパラギン酸であると

き、そのモチーフは、Uに選択的に結合でき；

(2-2) A_4 、 L_{ii} が、順に、アスパラギン、アスパラギンであるとき、そのモチーフは、Cに選択的に結合でき；

(2-3) A_4 、 L_{ii} が、順に、トレオニン、アスパラギンであるとき、そのモチーフは、Aに選択的に結合でき；

(2-4) A_4 、 L_{ii} が、順に、トレオニン、アスパラギン酸であるとき、そのモチーフは、Gに選択的に結合でき；

(2-5) A_4 、 L_{ii} が、順に、セリン、アスパラギンであるとき、そのモチーフは、Aに選択的に結合でき；

(2-6) A_4 、 L_{ii} が、順に、グリシン、アスパラギン酸であるとき、そのモチーフは、Gに選択的に結合でき；

(2-7) A_4 、 L_{ii} が、順に、アスパラギン、セリンであるとき、そのモチーフは、Cに選択的に結合でき；

(2-8) A_4 、 L_{ii} が、順に、プロリン、アスパラギン酸であるとき、そのモチーフは、Uに選択的に結合でき；

(2-9) A_4 、 L_{ii} が、順に、グリシン、アスパラギンであるとき、そのモチーフは、Aに選択的に結合でき；

(2-10) A_4 、 L_{ii} が、順に、メチオニン、アスパラギン酸であるとき、そのモチーフは、Uに選択的に結合でき；

(2-11) A_4 、 L_{ii} が、順に、ロイシン、アスパラギン酸であるとき、そのモチーフは、Cに選択的に結合でき；

(2-12) A_4 、 L_{ii} が、順に、バリン、トレオニンであるとき、そのモチーフは、Uに選択的に結合できる。

[請求項4] 請求項1に定義されたPPRモチーフを、1個以上（好ましくは2～14個）含む、RNA結合性蛋白質の標的となる塩基、又は塩基配列を同定する方法であって：

同定が、請求項2に記載の(3-1)～(3-16)のいずれか、又は請求項3に記載の(2-1)～(2-12)のいずれかに基づいて、PPRモチーフの A_1 、 A_4

、及び L_{ii} の3つのアミノ酸の組み合わせ、又は A_4 、 L_{ii} の2つのアミノ酸の組み合わせに応じた塩基の有無を認定することにより行われる、方法。

[請求項5] 標的RNA塩基、又は特定の塩基配列を有する標的RNAに結合可能な、請求項1に定義されたPPRモチーフを1個以上（好ましくは2～14個）含む、PPR蛋白質を同定する方法であって：

同定が、請求項2に記載の(3-1)～(3-16)のいずれか、又は請求項3に記載の(2-1)～(2-12)のいずれかに基づいて、標的RNA塩基、又は標的RNAを構成する特定の塩基に応じた、PPRモチーフの A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせの有無を認定することにより行われる、方法。

[請求項6] 請求項1に記載された方法で設計された蛋白質を用いる、RNAの機能の制御方法。

[請求項7] 請求項1に記載された方法で設計された蛋白質からなる領域と機能性領域とが連結されてなる、複合体。

[請求項8] 以下の工程を含む、細胞の遺伝物質を改変する方法：

標的配列を有するRNAを含む細胞を準備し；そして

請求項7に記載された複合体を細胞に導入することにより、複合体の蛋白質領域が標的配列を有するRNAに結合し、そのため機能性領域が、標的配列を有するRNAを改変する、方法。

[請求項9] 細胞質雄性不稔性の稔性回復因子として働くPPR蛋白質遺伝子において、様々な品種間で見られる当該遺伝子のアミノ酸多型を検出する工程；

当該遺伝子における多型と稔性との関連を特定する工程；

被検サンプルから得られたPPR蛋白質遺伝子の塩基配列を特定し、被検サンプルの稔性を決定する工程；

を含む、PPR蛋白質遺伝子の稔性を判定する方法。

[請求項10] PPR蛋白質が、式1で表される30～38アミノ酸長のポリペプチドから

なるPPRモチーフ

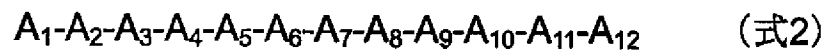
[化4]



(式中：

Helix Aは、12アミノ酸長の、 α ヘリックス構造を形成可能な部分であって、式2で表され、

[化5]



式2中、 $A_1 \sim A_{12}$ はそれぞれ独立にアミノ酸を表し；

Xは、存在しないか又は1～9アミノ酸長からなる部分であり；

Helix Bは、11～13アミノ酸長からなる、 α ヘリックス構造を形成可能な部分であり；

Lは、2～7アミノ酸長の、式3で表される部分であり；

[化6]



式3中、各アミノ酸は、“i” (-1)、“ii” (-2)、とC末端側からナンバリングされ、

ただし、 $L_{iii} \sim L_{vii}$ は存在しない場合がある。))

を1個以上（好ましくは2～16個）含む、蛋白質である、請求項9に記載の方法。

[請求項11] アミノ多型をPPRモチーフごとの多型として特定することを特徴とする、請求項9または10に記載の方法。

[請求項12] PPRモチーフの多型が、式1のモチーフの A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせ、又は A_4 、 L_{ii} の2つのアミノ酸の組み合わせによ

り特定されるものである、請求項9～11のいずれか1項に記載の方法。

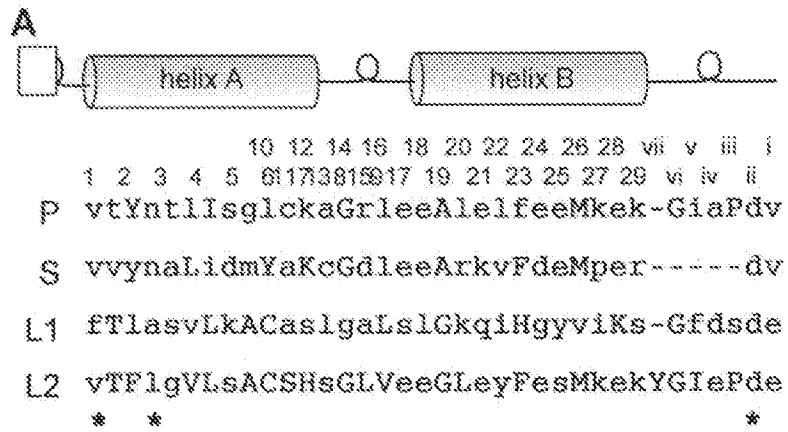
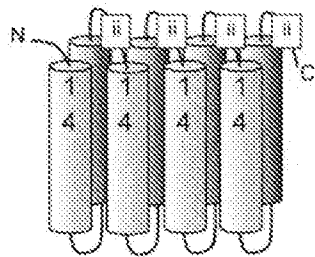
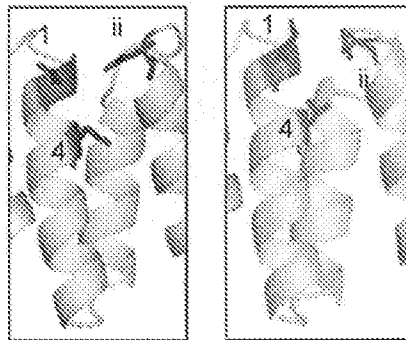
[請求項13] PPRモチーフの多型が、式1のモチーフの4番アミノ酸 (A_4) の多型により特定される、請求項12に記載の方法。

[請求項14] PPR蛋白質遺伝子上のすべてのPPRモチーフの4番アミノ酸が、Enko Bの対応するすべてのPPRモチーフの4番アミノ酸と同一であること、またはEnko Bの対応するすべてのPPRモチーフの“ii”番アミノ酸と同一であること、が、稔性であることを示す、請求項13に記載の方法。

[請求項15] PPR蛋白質遺伝子が、Enko Bをコードする「687遺伝子」と相同な遺伝子座に座乗するファミリー遺伝子、Enko Bと90%以上のアミノ酸配列同一性を有する遺伝子、Enko Bをコードする「687遺伝子」90%以上の塩基配列同一性を有する遺伝子である、請求項9～14のいずれか1項に記載の方法。

[請求項16] 様々な品種のorf687様遺伝子によりコードされる蛋白質が、配列番号：576～578、585～591のいずれかである、請求項9～15のいずれか1項に記載の方法。

[図1]

**B****C**

[図2]

Name	AGI number	Localization	PPH code	Target RNA sequence	Editing site	Cell*
AH011	A1g04880	-	FND-FTD-YTS-YTD-VST-ITH-VND-YLK-KTG-AND-IR-EGN	-	Unknown	-
CLB19	A1g09790	Chl	VTN-IF-LMTN-VNW-AND-VIN-FNT-VND-YTR-EGI	ccgagggcga cgagggccca	tpaA(78891) cpP(85542)	[1]
CRP21	A1g06740	Chl	EGN-YTN-FAG-FPG-FEN-VNT-VSD-ITD-VND-VAD-VSD-IRN-IRN-ITP-FTL-STE-PND-ITD-EGG	ccgagggcga cgagggccca	ndpD(118788)	[2]
CRP22	A1g11290	Chl	FTL-VHV-YTD-FTD-VNS-VL-VTN-VNT-VMN-SKT-VND-FVN-FTN-TNN-VLS-DGA	cgagggccca cgagggccca	ndpD(118788) ndpD(118788)	[3]
CRP28	A1g09720	Chl	FGS-FND-HPD-VNS-VND-YTD-LND-ANN-VVA-EGD	cgagggccca	ndpD(118788)	[3]
CRP34	A1g04890	Chl	FND-FED-FND-VNN-IND-IRN-TDD-VND-TVG-KVS-GRD-IVP-QGN	cgagggccca	ndpD(118788)	[4]
LO1	A1g14850	Mt	DGP-FNN-VTN-FPD-FCN-ENR-ICD-SNV-SAS-FST-FSN-VNN-MVG-EST	cgagggccca	ndpD(118788)	[8]
LPA86	A1g04870	Chl	LAM-FNN-FPD-FSE-VNS-VNN-VVD-VSN-SS-VIN-EGD	cgagggccca	ndpD(118788)	[8]
MEF1	A1g02100	Mt	ND-LNS-TSD-HPD-FEN-VSN-YSS-FSN-QNN-ILT-KAW	cgagggccca	ndpD(118788)	[7]
MEP9	A1g02060	Mt	VBN-TSD-PNV-ATA-VLL-VNS-VND-IRN-MSH-VND-ITD-FNE-IRN-IVD-ESD	cgagggccca	ndpD(118788)	[8]
MEP14	A1g02870	Mt	FNT-SPD-FST-SSD-MVG-SND-VNN-ISO-TTN-RTN-IVC-EGC	cgagggccca	ndpD(118788)	[8]
MEF18	A1g19020	Mt	RYN-VND-ANE-YTN-VAR-FTN-VND-VSS-VYV-FAH-AND-FND-IVN-PAK-SAN-VIC-KDD	cgagggccca	ndpD(118788)	[10]
MEP19	A1g02040	Mt	ISS-YND-FNN-YTN-VGN-TVN-ITT-VND-VLD-ACD-ITD-ILT-EGN	cgagggccca	ndpD(118788)	[10]
MEP20	A1g10970	Mt	RVS-LTN-ADT-YVD-FTN-FTN-IV-EGG	cgagggccca	ndpD(118788)	[10]
MEP21	A1g02040	Mt	FTN-FND-FNN-VND-IRN-ASD-IRD-GRN-IRN-ILK-EGD	cgagggccca	ndpD(118788)	[10]
MEP22	A1g12770	Mt	FTD-FND-FPG-FNT-YTD-VVE-IRN-IRN-ITD-FED-VSN-VLD-GAG	cgagggccca	ndpD(118788)	[10]
cp050	A1g03090	Chl	FFN-YTN-VSS-ILD-VTD-VTN-FVS-FND-IRN-VYD-EGG	cgagggccca	ndpD(118788)	[11]
cp051	A1g02970	Chl	FFN-YPD-FND-VNS-VVN-TND-VTN-ASN-IVN-HSD-FSN-VTE-KAS	cgagggccca	ndpD(118788)	[11]
cp082	A1g03070	Chl	LNN-YPD-YTY-AND-VND-SVN-KND-IRN-VLA-STE-SND-IRN-EGD	cgagggccca	ndpD(118788)	[11]
cp084	A1g02740	Chl	ED-YPS-TNN-VNS-FVN-FND-VND-FEN-FSW-AAN-TAD-FNV-TTN-MDA-ASN-IRN-VIS-DAN	cgagggccca	ndpD(118788)	[11]
cp085	A1g02880	Chl	VND-YND-YPC-VNN-ILY-KTS-ESD-ILS-KGL	cgagggccca	ndpD(118788)	[11]
cp058	A1g03370	Chl	FGY-FND-EPY-FND-VNN-YVE-YND-VND-VTN-QND-ITD-MOD-VND-YTD-VLE-SVD-LTD-ILW-EV7	cgagggccca	ndpD(118788)	[11]
cp057	A1g17480	Mt	FND-VNN-IRY-VNN-YND-VSD-FTS-VTN-CTD-SAR-NNQ-PCD-TSD-AAD-SAG-USD-VSD-FSE-SSD-ITD-YVE-IRN	cgagggccca	ndpD(118788)	[12]
SLO1	A1g02240	Mt	FND-FPV-HND-YND-VIT-FNT-VTN-AGD-IRN-ATN-LTD-IRN-EGD	cgagggccca	ndpD(118788)	[13]
Y81	A1g02880	Chl	FND-YPD-FNV-SCR-VNN-LSN-DND-ILW-NTT-VNW-TSD-VND-FTD-STD-VVE-VGN	cgagggccca	ndpD(118788)	[14]
PpPPH_58	-	Mt	FVN-HNS-ANN-GND-ITD-ITN-IVD-ITD-VND-ILD-RTN-VND-VVD-SND	cgagggccca	ndpD(118788)	[15]
PpPPH_71	-	Mt	RVN-YNT-VNS-ILD-ITD-SAR-ILD-ITD-VTD-IRN-CTD-VSD-VN-FND-IRN-VYT-KGD	cgagggccca	ndpD(118788)	[16]
PpPPH_77	-	Mt	AVD-FND-IRN-IRN-ND-VND-VND-VTD-VNN-TLD-DND-ING-VL-IL-ND-IRN-AD-ND-MTV-SED-AND-VNN-FVD-RAN-VND-STT-EGN	cgagggccca	ndpD(118788)	[16]
PpPPH_76	-	Mt	SND-YNT-HND-TMD-KNS-VTN-SD-ATD-IRN-IVD-SND-ITN-VTD-HND-VND-VSD-GRN-ND-VYT-EGN	cgagggccca	ndpD(118788)	[17]
PpPPH_79	-	Mt	YVN-YNS-VND-FVD-YND-VTE-IRN-ITD-IRN-ATL-FND-VTN-VND-AND-VNN-AYT-KAN	cgagggccca	ndpD(118788)	[17]
PpPPH_81	-	Mt	SCR-NDQ-7ND-VSN-ILD-ATN-VTN-VVD-IRN-MRN-VND-SIN-VTN-VND-VTD-VTN-VTN-EGD	cgagggccca	ndpD(118788)	[18]

[References]

- Chateignier-Boudin, A.L., et al., Plant J., 2008, 56:590-602.
- Okuda, K., et al., Proc. Natl. Acad. Sci. USA, 2007, 104:8178-8183.
- Okuda, K., et al., Plant Cell, 2009, 21:146-156.
- Kolera, E., M. Tasaka, and T. Shikama, Nature, 2005, 433 p. 323-330.
- Tang, J., et al., Plant J., 2010, 61:455-466.
- Cai, W., et al., Plant Physiol., 2009, 150:1250-1273.
- Zehrmann, A., et al., Plant Cell, 2009, 21:558-567.
- Takenaka, M., Plant Physiol., 2009, 152:939-947.
- Verbitsky, D., et al., FEBS Lett., 2011, 585:700-704.
- Takenaka, M., et al., J. Biol. Chem., 2010, 285:27122-27129.
- Hammani, K., et al., Plant Cell, 2009, 21:3686-3699.
- Hammani, K., et al., J. Biol. Chem., 2011, 286:21261-21271.
- Gung, T.-Y., et al., Plant J., 2010, 63:499-511.
- Zhou, W., et al., Plant J., 2009, 58:82-93.
- Ohtani, S., et al., Plant Cell Physiol., 2010, 51:1942-1949.
- Tasaki, E., et al., Plant J., 2010, 62:560-570.
- Uchida, M., et al., FEBS Lett., 2011, 585:2367-2371.

[図3-1]

モザ…ア番号

XZ01																																				
アZ/総番号																																				
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34			
1	P	2	N	N	V	1	1	G	P	2	E	S	R	9	P	K	E	2	F	L	L	F	K	2	M	L	R	M	2	D	G	E	1	R		
2	F	1	F	2	V	1	F	A	V	2	A	D	L	K	2	S	L	G	H	M	1	1	G	M	V	L	A	L	2	E	E	L	F	1		
3	H	V	N	N	A	S	1	1	M	1	A	S	2	2	2	N	A	N	A	2	V	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
4	V	2	N	N	C	L	1	N	G	1	K	K	1	G	E	A	E	S	2	1	V	1	K	L	M	F	S	E	2	V	R	D	1	U		
5	V	1	M	1	2	L	V	A	S	2	2	N	L	G	2	L	N	M	U	K	E	F	1	E	V	1	F	S	N	2	G	L	R	M	1	
6	P	1	1	N	A	L	M	2	M	1	S	E	2	G	2	1	H	E	A	P	D	1	P	G	H	L	E	R	A	2	1	A	R	1		
7	V	S	K	1	1	M	1	S	G	V	A	R	G	2	L	L	2	V	A	R	V	L	P	2	2	2	2	2	2	2	2	2	2	2	2	2
8	A	N	1	2	G	1	V	2	A	E	2	2	G	2	A	L	A	L	2	2	E	N	2	2	N	1	X	P	2	2	2	2	2	2	2	
9	V	1	M	1	N	2	1	2	A	G	S	G	L	N	A	L	D	V	2	1	M	1	1	P	1	1	E	S	V	2	L	N	2	1	V	
10	A	1	S	Y	2	1	V	1	M	1	K	F	U	S	N	1	S	E	A	1	2	V	1	N	G	L	R	1	2	1	A	R	1	2	1	
11	L	1	Y	1	A	1	1	1	G	L	2	L	H	G	2	A	S	T	A	1	S	1	P	N	E	M	1	D	A	2	1	A	R	1		
12	1	1	F	1	2	L	L	1	1	A	G	2	G	H	2	G	2	1	2	1	G	R	D	V	1	2	2	2	2	2	2	2	2	2	2	
13	K	K	V	2	1	N	V	2	L	L	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
14	K	H	V	2	1	M	V	2	L	L	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	

MSF1

モザ…ア番号

アZ/総番号																																					
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34				
1	N	G	1	2	D	L	L	1	S	2	A	R	T	P	2	1	F	G	1	G	L	H	2	V	V	V	K	S	2	G	L	R	L	1	P		
2	L	1	A	N	N	1	1	1	A	G	S	G	L	P	F	U	S	P	A	E	2	S	P	U	K	2	2	2	2	2	2	2	2	2	2		
3	T	1	N	S	2	1	1	1	G	P	A	G	H	E	1	P	W	M	S	L	E	F	1	K	1	M	M	A	G	2	N	L	R	P	1	U	
4	N	V	L	P	S	A	1	K	S	C	A	1	L	S	2	2	U	1	2	N	S	V	1	2	1	S	M	K	1	2	Y	G	A	1	U		
5	F	1	U	S	L	1	V	2	M	1	A	K	2	G	E	1	V	1	A	P	K	M	1	D	2	M	1	G	R	2	2	2	2	2	2		
6	V	1	N	2	G	N	M	1	G	V	A	G	M	2	E	N	E	E	A	L	2	P	R	E	A	L	P	2	2	N	L	A	V	2	2		
7	V	S	F	S	V	1	2	1	L	L	L	L	M	U	1	H	2	L	S	1	K	2	2	2	2	2	2	2	2	2	2	2	2	2	2		
8	F	1	G	S	S	L	V	A	L	1	2	2	K	G	V	P	E	M	A	V	1	N	V	1	M	2	V	2	V	N	2	2	2	2	2		
9	G	1	W	R	A	M	1	L	K	A	V	E	P	N	K	U	1	K	E	A	V	D	L	P	1	1	M	N	1	E	2	G	E	K	P	1	
10	1	1	F	1	N	V	1	L	N	A	G	S	N	A	G	L	V	2	G	E	2	N	1	Y	F	2	2	2	M	K	E	S	2	1	L	P	1
11	N	H	1	A	S	L	V	2	M	L	G	R	A	G	2	L	G	E	A	L	E	V	1	1	N	M	1	L	D	2	2	2	2	2	2	2	

MSF2

モザ…ア番号

アZ/総番号																																			
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34		
1	V	1	N	N	T	M	1	1	G	V	V	K	N	E	M	2	G	A	R	V	L	P	2	2	2	2	2	2	2	2	2	2	2	2	2
2	F	M	1	S	2	1	V	2	U	G	2	1	R	P	L	E	E	A	P	K	L	P	D	E	M	P	2	S	2	2	2	2	2	2	2
3	F	S	M	K	1	M	1	2	2	V	A	K	N	R	K	1	S	E	A	L	1	L	P	E	K	M	P	E	R	2	N	A	V	S	N
4	A	N	1	1	G	1	2	G	2	E	V	D	S	A	V	1	L	P	R	M	P	1	2	D	N	S	2	2	2	2	2	2	2	2	2
5	V	A	S	1	1	K	N	E	N	L	S	E	A	M	1	L	G	2	2	S	L	V	S	G	2	E	D	2	2	2	2	2	2	2	2
6	V	A	1	2	1	1	1	1	2	V	G	2	R	G	2	V	E	A	A	H	1	L	P	D	2	1	2	D	L	2	2	2	2	2	2
7	V	S	N	N	1	1	1	1	A	V	L	P	V	2	2	V	S	A	R	L	L	P	G	N	K	G	2	2	2	2	2	2	2	2	2
8	1	L	N	N	T	M	1	U	G	V	V	N	V	1	S	M	E	S	A	F	A	L	P	S	2	2	M	P	N	R	2	A	H	S	N
9	M	N	V	S	V	1	2	2	V	G	2	V	E	L	A	R	M	V	E	E	A	T	1	E	K	2	2	2	2	2	2	2	2	2	2
10	V	S	M	N	C	1	1	A	A	V	E	P	N	K	U	1	K	E	A	V	D	L	P	1	1	M	N	1	E	2	G	E	K	P	1
11	N	1	L	Y	S	L	1	L	A	2	2	2	L	V	N	1	N	1	2	M	2	M	H	2	1	V	V	K	1	2	2	2	2	2	2
12	P	V	1	N	A	L	1	1	M	V	2	N	2	G	E	1	M	E	2	R	R	1	F	D	2	2	M	L	K	2	2	2	2	2	2
13	1	1	W	R	A	M	1	G	2	V	A	P	H	R	N	A	S	E	A	L	H	L	P	G	2	M	K	S	R	2	1	Y	P	1	U
14	1	1	F	1	V	2	S	2	L	N	A	O	A	M	A	2	L	V	2	E	A	K	A	G	P	V	2	M	M	S	V	C	K	1	E
15	E	N	1	S	S	L	V	N	1	V	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2

MSF15

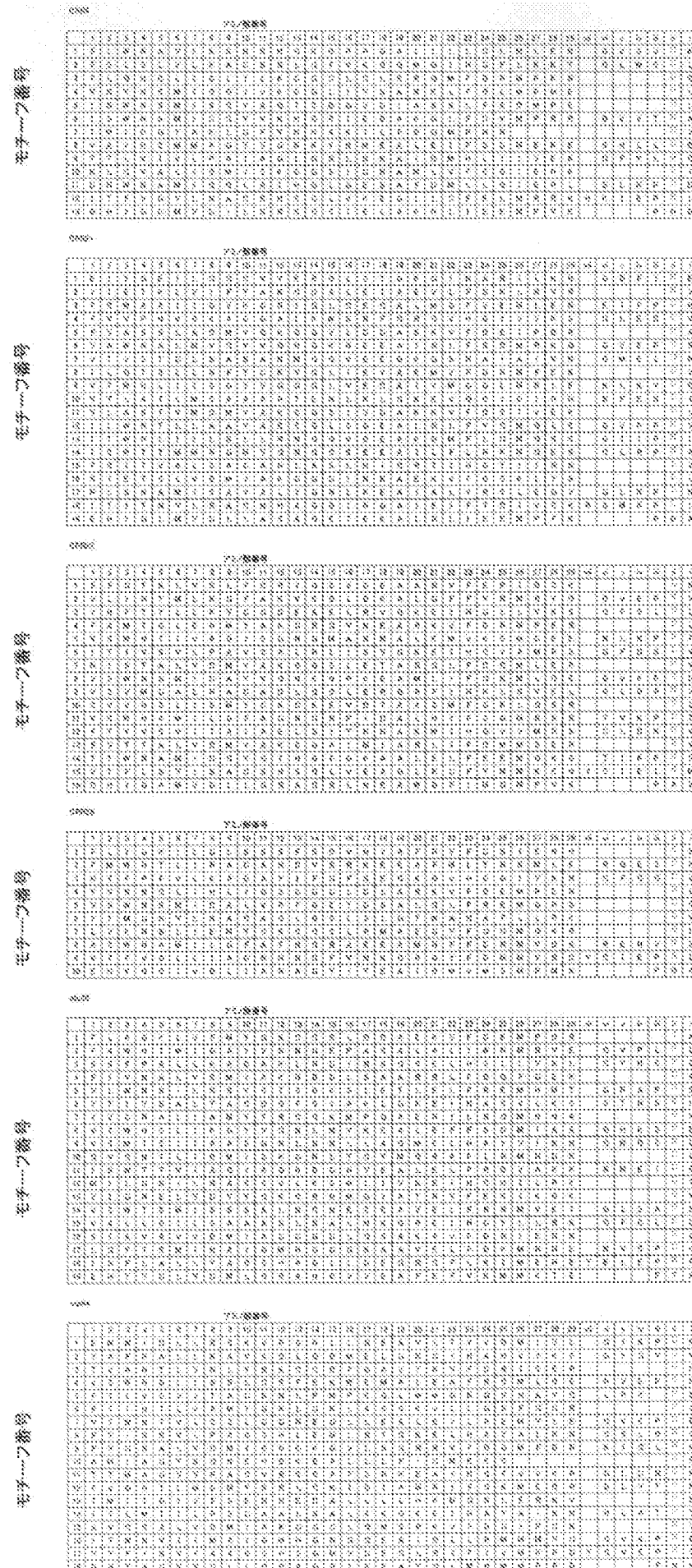
モザ…ア番号

アZ/総番号																																				
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34			
1	P	A	L	V	2	A	L	2	2	A	2	S	N	D	1	1	2	2	N	G	1	M	2	P	1	L	K	S	2	G	L	D	S	1	G	
2	V	1	L	N	S	1	1	1	M	1	V	A	K	U	N	L	L	A	S	A	S	2	V	1	R	2	M	P	K	L	2	2	2	2	2	
3	A	G	E	R	1	M	V	D	2	V	2	S	N	L	M	2	2	A	L	K	1	V	2	M	2	D	N	S	2	2	2	2	2	2	2	
4	V	S	Y	2	1	1	1	1	P	G	V	A	G	M	2	E	N	E	E	A	M	1	L	P	E	K	M	P	E	R	2	N	A	V	S	N
5	V	1	L	K	1	V	1	1	S	A	G	N	L	G	2	1	M	2	N	M	1	L	G	S	L	A	1	K	L	2	1	M	E	G	A	Y
6	P	V	1	N	2	G	N	M	1	G	V	A	G	M	2	E	N	E	E	A	L	2	P	R	E	A	L	P	2	2	N	L	A	V	2	2
7	V	1	N	N	V	M	1	N	G	V	S	P	A	S	L	1	L	E	S	A	K	E	2	P	D	G	2	2	E	R	2	2	2	2	2	
8	V	S	M	N	C	1	1	A	A	V	E	P	N	K	U	1	K	E	A	V	D	L	P	1	1	M	N									

44-38861-107

[illegible]

[図3-3]



• **Stress**

1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	
2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100		
3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100			
4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100				
5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94											

下元ノ節候表																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																								
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	

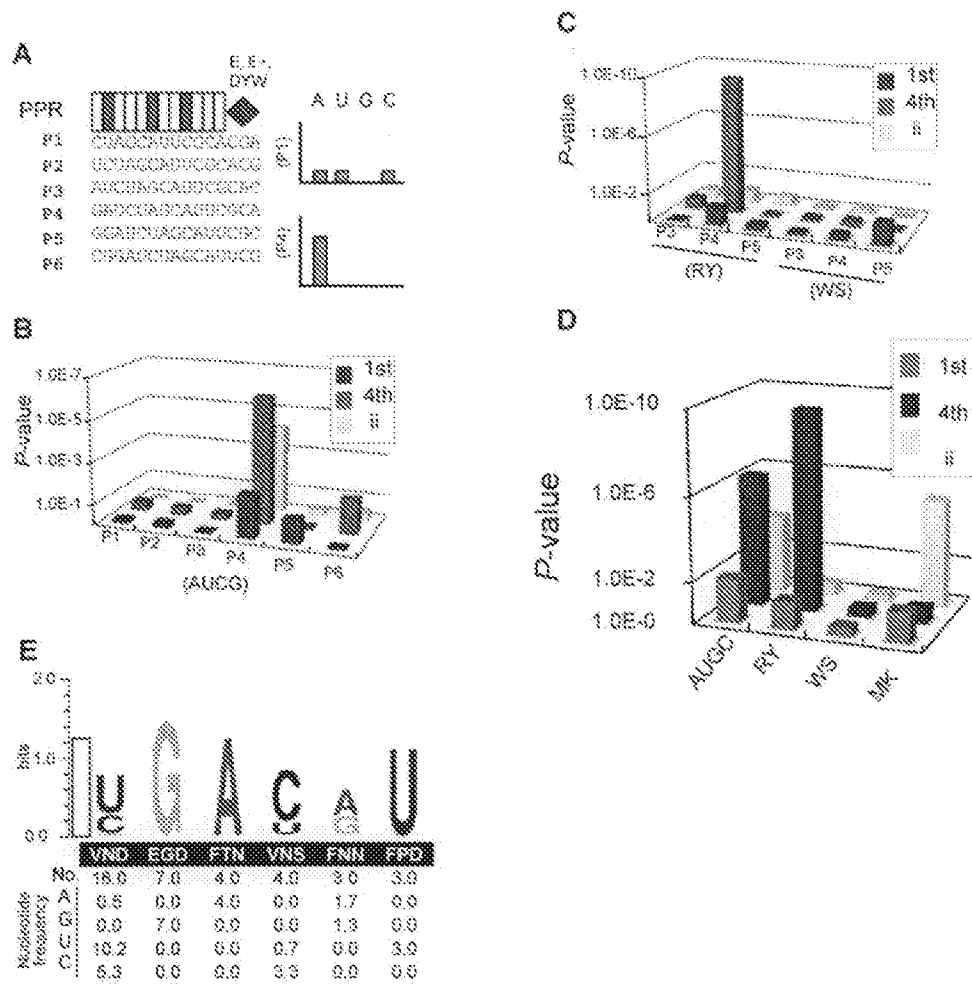
[illegible]

2022		2021		2020		2019		2018		2017		2016		2015		2014		2013		2012		2011		2010		2009		2008		2007		2006		2005		2004		2003		2002		2001		2000		1999		1998		1997		1996		1995		1994		1993		1992		1991		1990		1989		1988		1987		1986		1985		1984		1983		1982		1981		1980		1979		1978		1977		1976		1975		1974		1973		1972		1971		1970		1969		1968		1967		1966		1965		1964		1963		1962		1961		1960		1959		1958		1957		1956		1955		1954		1953		1952		1951		1950		1949		1948		1947		1946		1945		1944		1943		1942		1941		1940		1939		1938		1937		1936		1935		1934		1933		1932		1931		1930		1929		1928		1927		1926		1925		1924		1923		1922		1921		1920		1919		1918		1917		1916		1915		1914		1913		1912		1911		1910		1909		1908		1907		1906		1905		1904		1903		1902		1901		1900		1899		1898		1897		1896		1895		1894		1893		1892		1891		1890		1889		1888		1887		1886		1885		1884		1883		1882		1881		1880		1879		1878		1877		1876		1875		1874		1873		1872		1871		1870		1869		1868		1867		1866		1865		1864		1863		1862		1861		1860		1859		1858		1857		1856		1855		1854		1853		1852		1851		1850		1849		1848		1847		1846		1845		1844		1843		1842		1841		1840		1839		1838		1837		1836		1835		1834		1833		1832		1831		1830		1829		1828		1827		1826		1825		1824		1823		1822		1821		1820		1819		1818		1817		1816		1815		1814		1813		1812		1811		1810		1809		1808		1807		1806		1805		1804		1803		1802		1801		1800		1799		1798		1797		1796		1795		1794		1793		1792		1791		1790		1789		1788		1787		1786		1785		1784		1783		1782		1781		1780		1779		1778		1777		1776		1775		1774		1773		1772		1771		1770		1769		1768		1767		1766		1765		1764		1763		1762		1761		1760		1759		1758		1757		1756		1755		1754		1753		1752		1751		1750		1749		1748		1747		1746		1745		1744		1743		1742		1741		1740		1739		1738		1737		1736		1735		1734		1733		1732		1731		1730		1729		1728		1727		1726		1725		1724		1723		1722		1721		1720		1719		1718		1717		1716		1715		1714		1713		1712		1711		1710		1709	
------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--	------	--

00000	P2 / 骰子号																																								
1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36					
2	1	2	1	Y	2	3	1	1	5	4	5	4	5	5	5	5	4	5	4	5	4	5	4	5	4	5	4	5	4	5	4	5	4	5	4	5	4				
3	2	3	4	5	6	1	2	3	4	5	6	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38				
4	3	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36				
5	4	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37				
6	5	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38				
7	6	4	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38			
8	7	5	4	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38		
9	8	6	5	4	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	
10	9	7	6	5	4	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38
11	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38			
12	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39			
13	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40			
14	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41			
15	5	6	7	8	9	10	11	12	13	14																															

500P4		502 / 總卷第																																																																																																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	
1	2	3	4	5	6</																																																																																															

[図4]



[図5]

整列P4、4番アミノ酸

実測値

	計	A	U	G	C
A	16.00	4.33	5.50	3.00	3.17
C	5.99	4.00	0.33	1.33	0.33
D	1.00	0.00	0.00	1.00	0.00
F	1.00	0.00	1.00	0.00	0.00
G	25.00	8.00	6.83	8.00	2.17
H	1.00	0.67	0.33	0.00	0.00
I	10.00	1.83	2.83	0.83	4.50
K	2.00	0.50	0.00	1.00	0.50
L	15.00	2.33	8.33	0.00	4.33
N	98.96	10.66	41.66	11.33	35.32
M	3.00	0.00	2.67	0.00	0.33
P	17.00	1.00	13.83	1.67	0.50
Q	1.00	0.00	1.00	0.00	0.00
R	1.00	1.00	0.00	0.00	0.00
S	41.99	15.83	12.33	8.83	5.00
T	53.99	26.67	7.17	16.17	4.00
V	33.00	9.50	13.17	2.83	7.50
計	326.92	86.32	116.98	55.98	67.64

P値

5.126E-07

理論値(中央値)

	計	A	U	G	C
A	16.00	4.22	5.72	2.74	3.31
C	5.99	1.58	2.14	1.03	1.24
D	1.00	0.26	0.36	0.17	0.21
F	1.00	0.26	0.36	0.17	0.21
G	25.00	6.60	8.94	4.28	5.17
H	1.00	0.26	0.36	0.17	0.21
I	10.00	2.64	3.58	1.71	2.07
K	2.00	0.53	0.72	0.34	0.41
L	15.00	3.96	5.37	2.57	3.10
N	98.96	26.12	35.40	16.94	20.47
M	3.00	0.79	1.07	0.51	0.62
P	17.00	4.49	6.08	2.91	3.52
Q	1.00	0.26	0.36	0.17	0.21
R	1.00	0.26	0.36	0.17	0.21
S	41.99	11.08	15.02	7.19	8.69
T	53.99	14.25	19.32	9.24	11.17
V	33.00	8.71	11.80	5.65	6.83
計	326.92	86.30	116.95	55.97	67.63

整列P4、5番アミノ酸

実測値

	計	A	U	G	C
A	73.979	13.165	28.658	15.662	16.494
C	26.995	7.666	9.499	8	1.83
D	3.999	2.333	0.333	1.333	0
E	4	1	2	0	1
F	6	0	4.5	0	1.5
G	34.984	8.329	12.331	6.163	8.161
H	2	0.5	1	0	0.5
I	5	1	1	0.5	2.5
K	6.999	2.666	3.333	0	1
L	5	1	1	1	2
N	7.999	4.333	2	0	1.666
S	94.974	28.494	34.992	10.995	20.493
T	40.995	11.332	11.665	9.166	8.832
V	10.999	2.5	4.333	2.5	1.666
Y	2.995	1.999	0.333	0.663	0
計	326.92	86.317	116.98	55.982	67.642

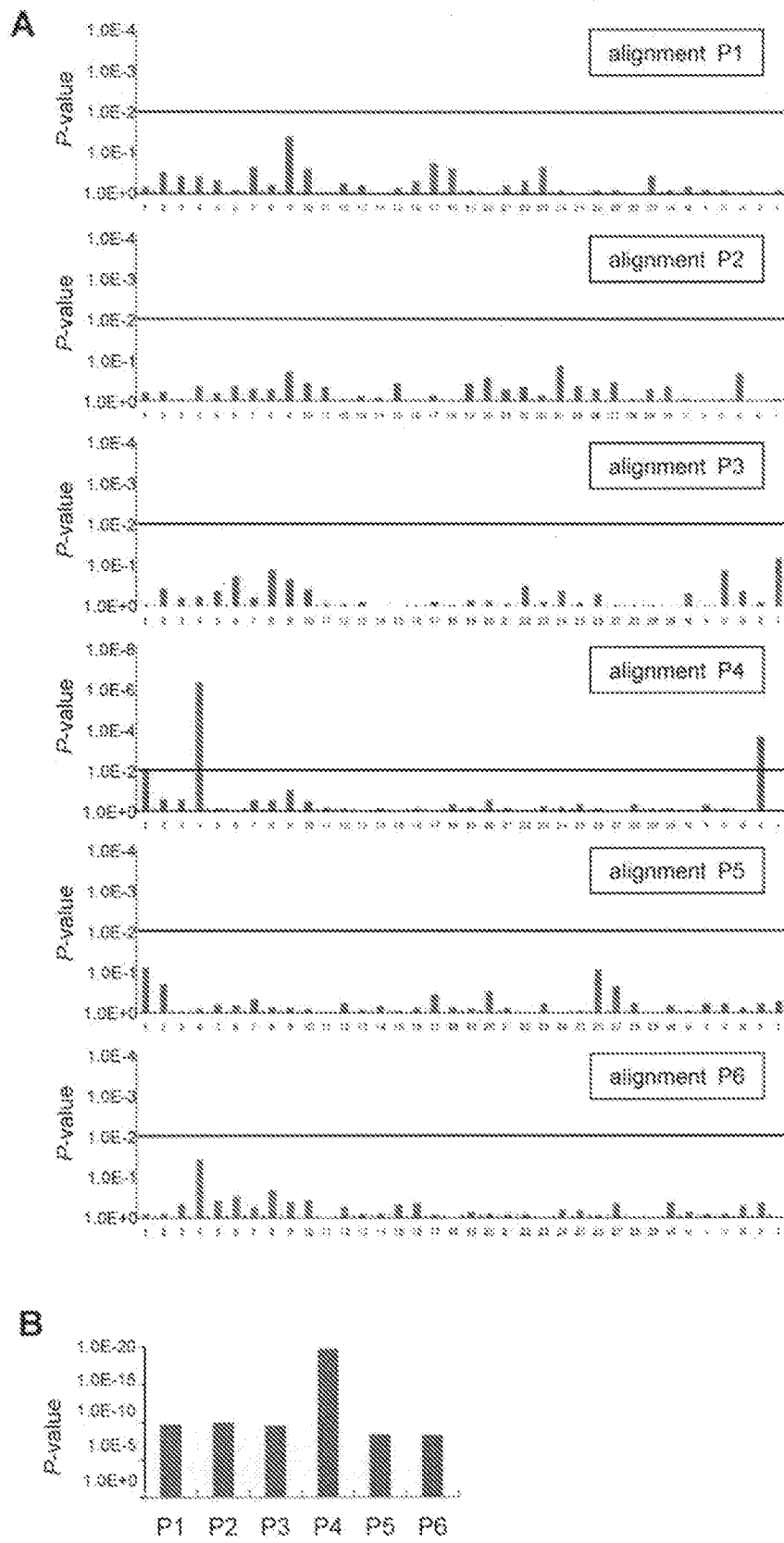
P値

0.739419764

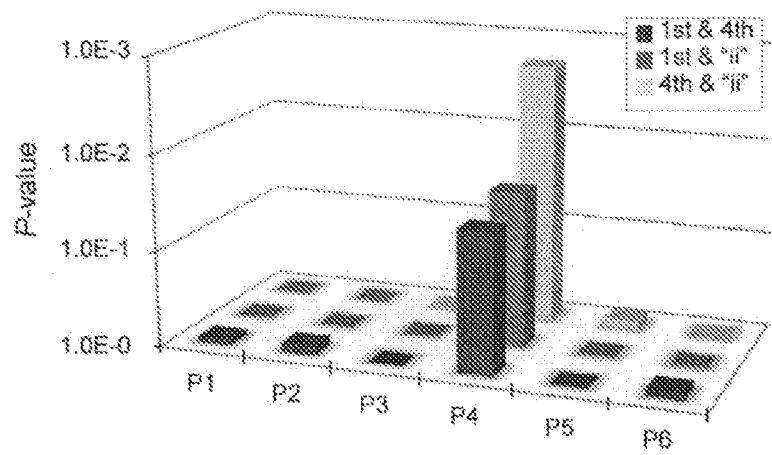
理論値(中央値)

	計	A	U	G	C
A	73.979	19.528	26.464	12.665	15.303
C	26.995	7.1258	9.6569	4.6215	5.5841
D	3.999	1.0556	1.4306	0.6846	0.8272
E	4	1.0559	1.4309	0.6848	0.8274
F	6	1.5838	2.1464	1.0272	1.2411
G	34.984	9.2346	12.515	5.9892	7.2367
H	2	0.5279	0.7155	0.3424	0.4137
I	5	1.3198	1.7886	0.856	1.0343
K	6.999	1.8475	2.5037	1.1962	1.4478
L	5	1.3198	1.7886	0.856	1.0343
N	7.999	2.1115	2.8615	1.3694	1.6546
S	94.974	25.07	33.975	16.259	19.646
T	40.995	10.821	14.665	7.0183	8.4801
V	10.999	2.9034	3.9346	1.883	2.2752
Y	2.995	0.7906	1.0714	0.5127	0.6195
計	326.92	86.30	116.95	55.97	67.63

[図6]



[図7]



[図8]

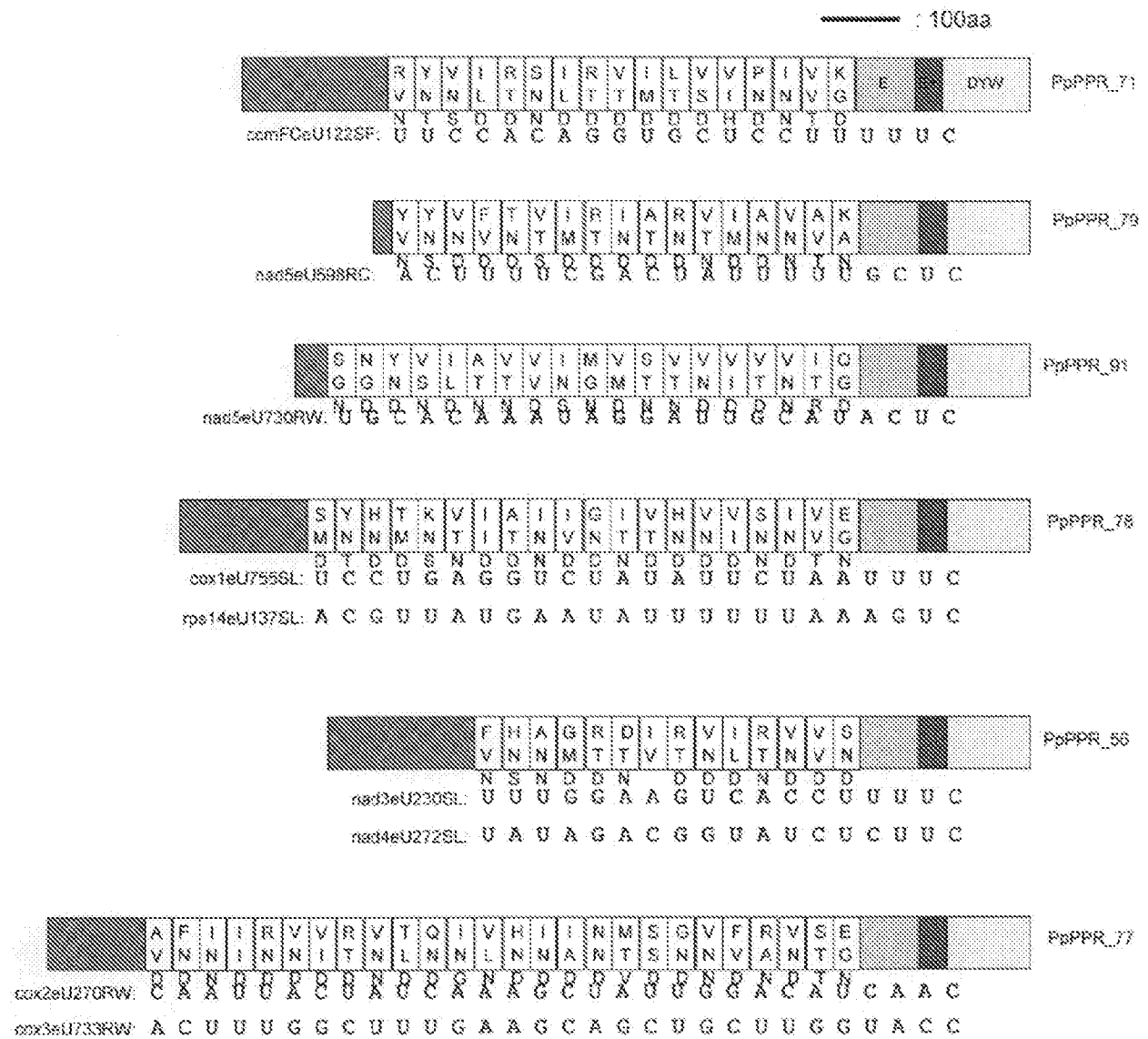
PPR code	occurrence of the code	Nucleotide occurrence frequency				*Scoring matrix				*Probability matrix (for MEME)			
		A	G	U	C	A	G	U	C	A	G	U	C
VND	18	0.50	0.00	10.17	5.33	-3.74	-2.75	4.42	2.01	0.00	0.00	0.84	0.33
FND	12	1.83	1.00	8.83	2.33	-1.34	-1.06	2.53	-0.16	0.15	0.08	0.87	0.19
EGD	7	0.00	7.00	0.00	0.00	-1.85	5.80	-2.51	-1.45	0.00	1.00	0.00	0.00
VNV	7	1.33	0.00	0.33	5.33	-0.52	-1.20	-2.18	3.88	0.19	0.00	0.05	0.79
FSN	8	4.67	0.33	0.33	0.67	3.08	-0.70	-1.82	-0.58	0.78	0.08	0.06	0.11
VTN	8	3.50	1.50	0.00	1.00	1.91	0.47	-2.15	-0.24	0.58	0.25	0.00	0.17
INN	5	0.00	0.00	1.83	3.17	-1.32	-0.88	0.04	2.13	0.00	0.00	0.37	0.63
ITD	5	0.00	3.00	1.00	1.00	-1.32	2.14	-0.79	-0.04	0.00	0.80	0.20	0.20
FTN	4	4.83	0.00	0.00	0.00	2.94	-0.89	-1.34	-0.83	1.00	0.00	0.00	0.00
TSD	4	0.33	0.50	2.50	0.67	-0.73	-0.19	1.06	-0.16	0.05	0.13	0.63	0.17
VNS	4	0.00	0.00	0.67	3.33	-1.06	-0.09	-0.77	2.50	0.00	0.00	0.17	0.83
VSN	4	2.67	0.00	1.00	0.33	1.61	-0.08	-0.44	-0.50	0.67	0.00	0.25	0.08
VND	4	0.00	0.50	3.00	0.50	-1.06	-0.19	1.58	-0.53	0.00	0.13	0.75	0.13
FNV	3	1.67	1.33	0.00	0.00	0.87	0.81	-1.08	-0.62	0.55	0.44	0.00	0.00
FNW	3	0.33	1.00	0.67	1.00	-0.46	0.48	-0.41	0.38	0.11	0.33	0.22	0.33
FPO	3	0.00	0.00	3.00	0.00	-0.79	-0.52	1.92	-0.52	0.00	0.00	1.00	0.00
INW	3	0.00	0.00	2.00	1.00	-0.79	-0.52	0.92	0.38	0.00	0.00	0.67	0.33
VNW	3	0.50	0.00	1.00	1.50	-0.29	-0.52	-0.08	0.86	0.17	0.00	0.33	0.50
VSD	3	1.50	1.00	0.50	0.00	0.71	0.48	-0.58	-0.62	0.50	0.33	0.17	0.00
STD	3	0.00	3.00	0.00	0.00	-0.79	2.45	-1.08	-0.62	0.00	1.00	0.00	0.00
VVE	3	1.00	0.00	1.50	0.50	0.21	-0.52	0.42	-0.12	0.33	0.00	0.50	0.17
YPO	3	0.00	0.00	3.00	0.00	-0.79	-0.52	1.92	-0.52	0.00	0.00	1.00	0.00
VTN	3	0.00	1.00	2.00	0.00	-0.79	0.45	0.92	-0.52	0.00	0.33	0.67	0.00
ND	48	3.83	5.00	30.50	8.67	2.84	-3.22	13.32	-1.27	0.08	0.10	0.64	0.18
NN	23	2.33	3.50	4.16	11.49	-0.74	-0.44	-4.07	7.24	0.14	0.15	0.18	0.52
TN	20	12.83	2.83	2.67	1.67	7.55	-0.59	-4.19	-2.47	0.64	0.14	0.13	0.08
SD	19	4.83	8.50	5.00	2.67	-0.16	3.25	-1.80	-1.27	0.25	0.34	0.26	0.14
TD	17	2.33	11.67	1.67	1.33	-2.16	8.75	-4.42	-2.18	0.14	0.69	0.10	0.08
SN	14	9.33	0.33	3.33	1.00	8.84	-2.06	-1.88	-1.90	0.67	0.02	0.24	0.07
GD	10	0.50	8.00	1.50	0.00	-2.14	6.29	-2.08	-2.07	0.06	0.80	0.15	0.00
NW	10	1.15	1.33	3.67	3.83	-1.48	-0.38	0.00	1.78	0.12	0.13	0.37	0.38
PD	9	0.00	0.67	7.83	0.50	-2.38	-0.88	4.81	-1.36	0.00	0.07	0.87	0.06
VN	9	2.67	1.33	3.00	2.00	0.29	-0.21	-0.22	0.14	0.30	0.15	0.33	0.22
NS	8	0.33	0.00	2.67	5.00	-1.78	-1.37	-0.20	3.34	0.04	0.00	0.33	0.62
NT	8	1.00	1.50	0.67	4.83	-1.11	0.13	-2.20	3.18	0.13	0.19	0.06	0.60
GN	6	3.00	0.00	2.00	0.00	1.68	-0.88	0.21	-1.03	0.60	0.00	0.40	0.00
VG	5	1.00	0.00	1.00	3.00	-0.32	-0.88	-0.79	1.97	0.20	0.00	0.20	0.60
AG	4	0.33	3.00	0.67	0.00	-0.72	2.39	-0.77	-0.83	0.08	0.75	0.17	0.00
SS	4	0.00	0.33	2.67	1.00	-1.06	-0.35	1.23	0.17	0.00	0.08	0.67	0.25
VE	4	1.00	0.00	2.50	0.50	-0.08	-0.58	1.07	-0.33	0.25	0.00	0.66	0.13
AN	3	0.67	0.00	1.67	0.67	-0.13	-0.51	0.59	0.05	0.22	0.00	0.56	0.22
CD	3	1.00	1.33	0.33	0.33	0.21	0.82	-0.74	-0.29	0.33	0.44	0.11	0.11
IN	3	0.50	0.00	1.50	1.00	-0.29	-0.51	0.45	0.38	0.17	0.00	0.50	0.33
TW	3	1.50	0.50	1.00	0.00	0.71	-0.01	-0.07	-0.62	0.50	0.17	0.33	0.00
VD	3	0.00	1.00	1.00	1.00	-0.79	0.48	-0.07	0.38	0.00	0.33	0.33	0.33
VS	3	1.00	0.00	1.00	1.00	0.21	-0.51	-0.07	0.38	0.33	0.00	0.33	0.33
*Background frequency		0.28	0.17	0.36	0.21	0.00	0.00	0.00	0.00	0.28	0.17	0.36	0.21

*The value indicates a probability of nucleotide occurrence for all RNA sequences surrounding the RNA editing sites used to construct the PPR code.

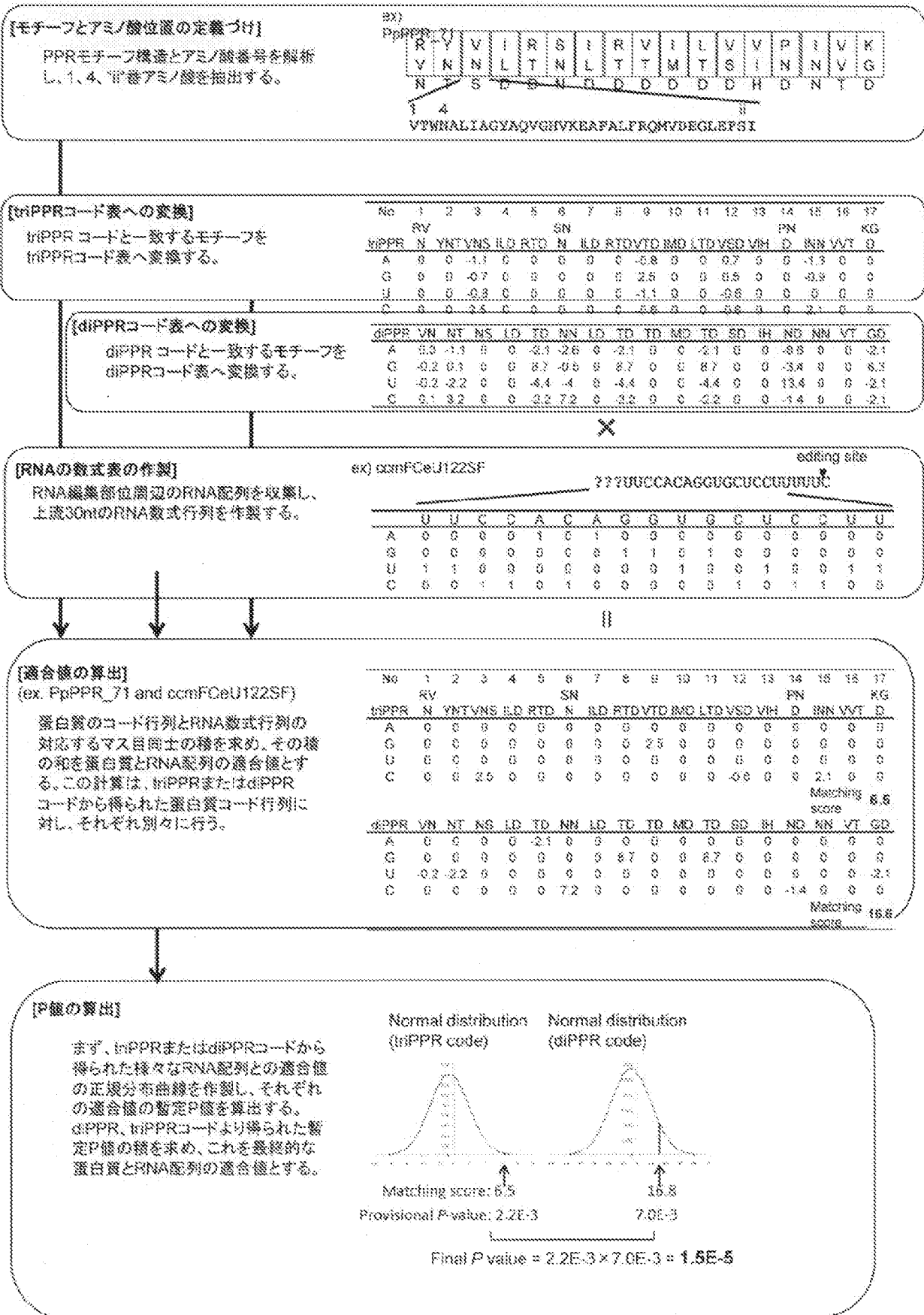
*The scoring matrix is estimated by a subtraction of the background frequency from the nucleotide occurrence frequency of the PPR code.

*The probability matrix is obtained from the nucleotide occurrence frequency.

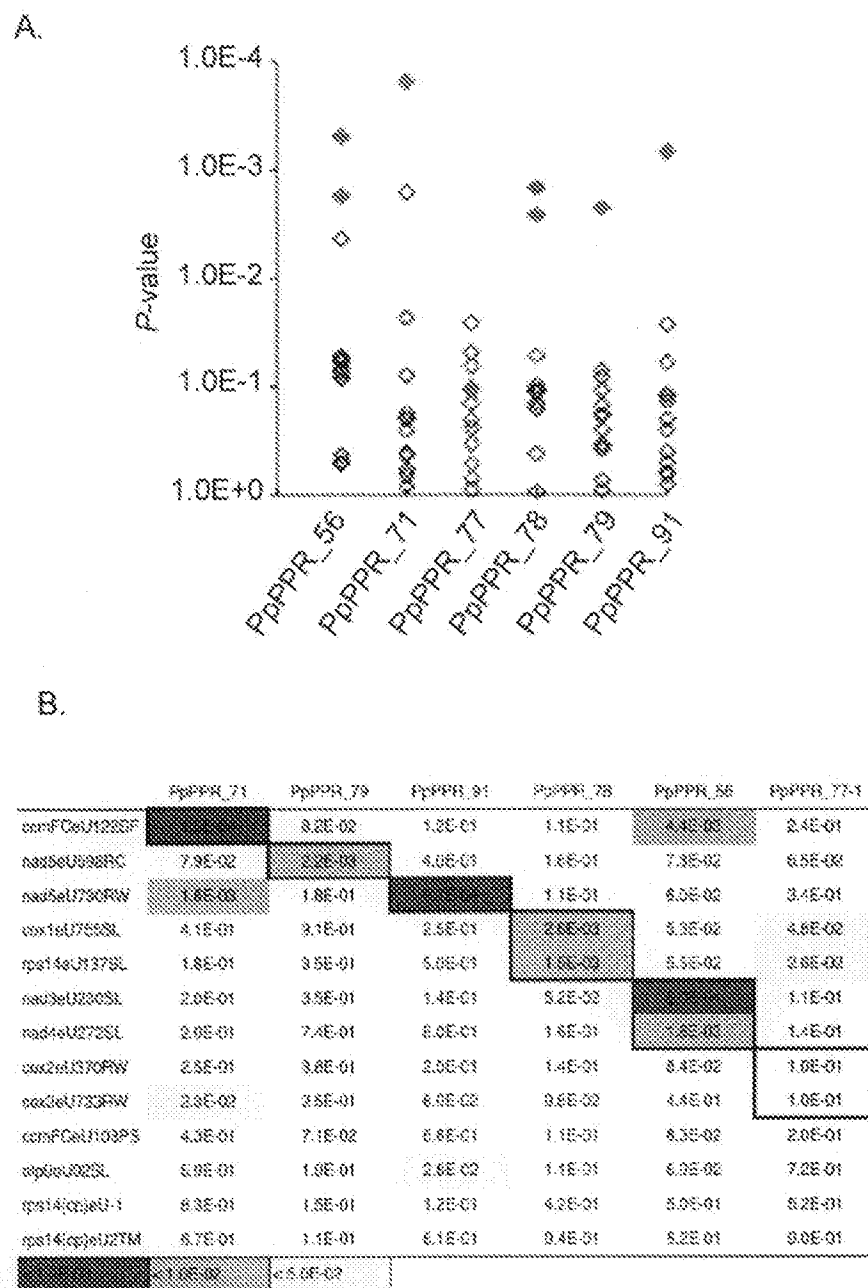
[図9]



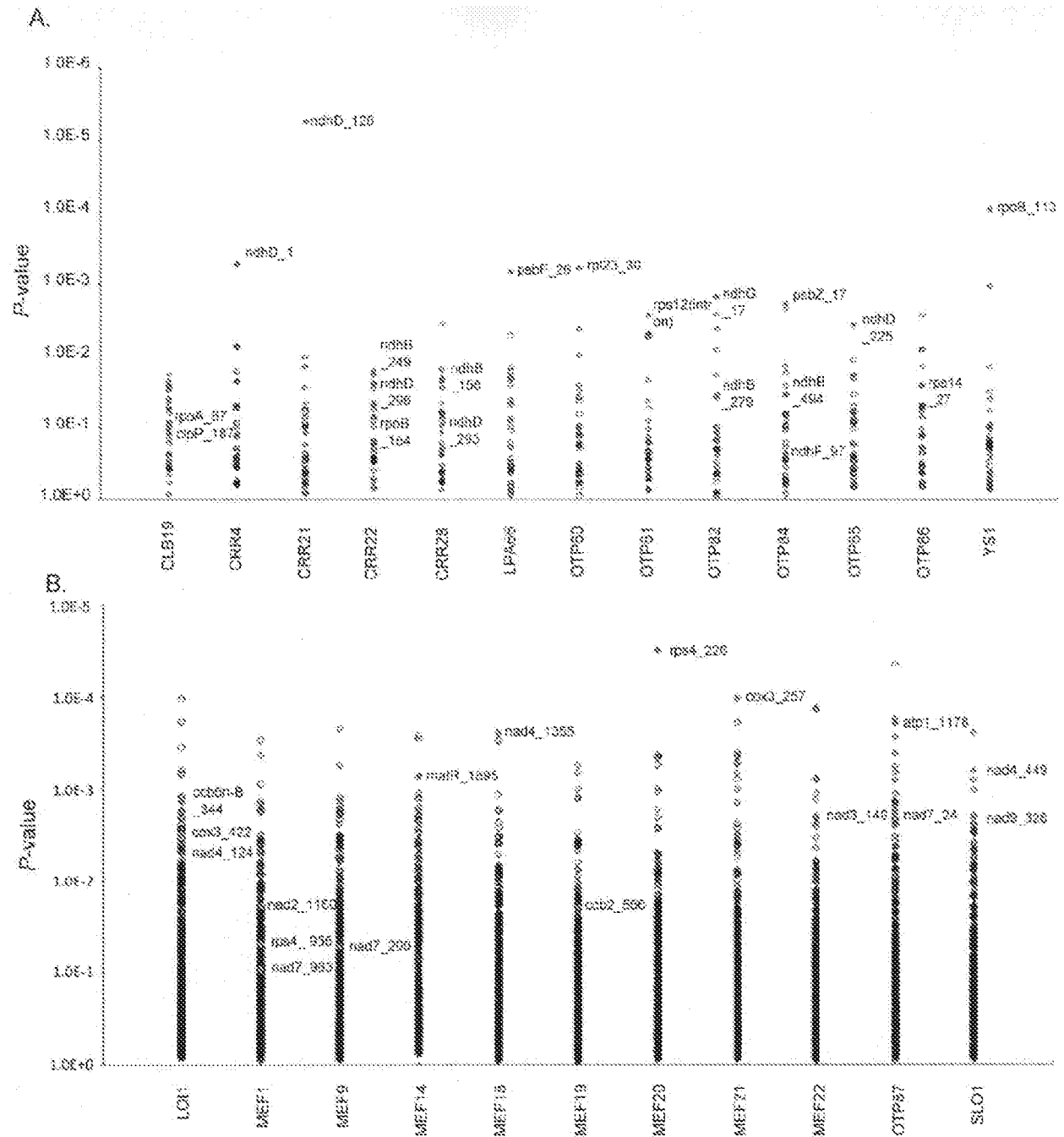
[図10]



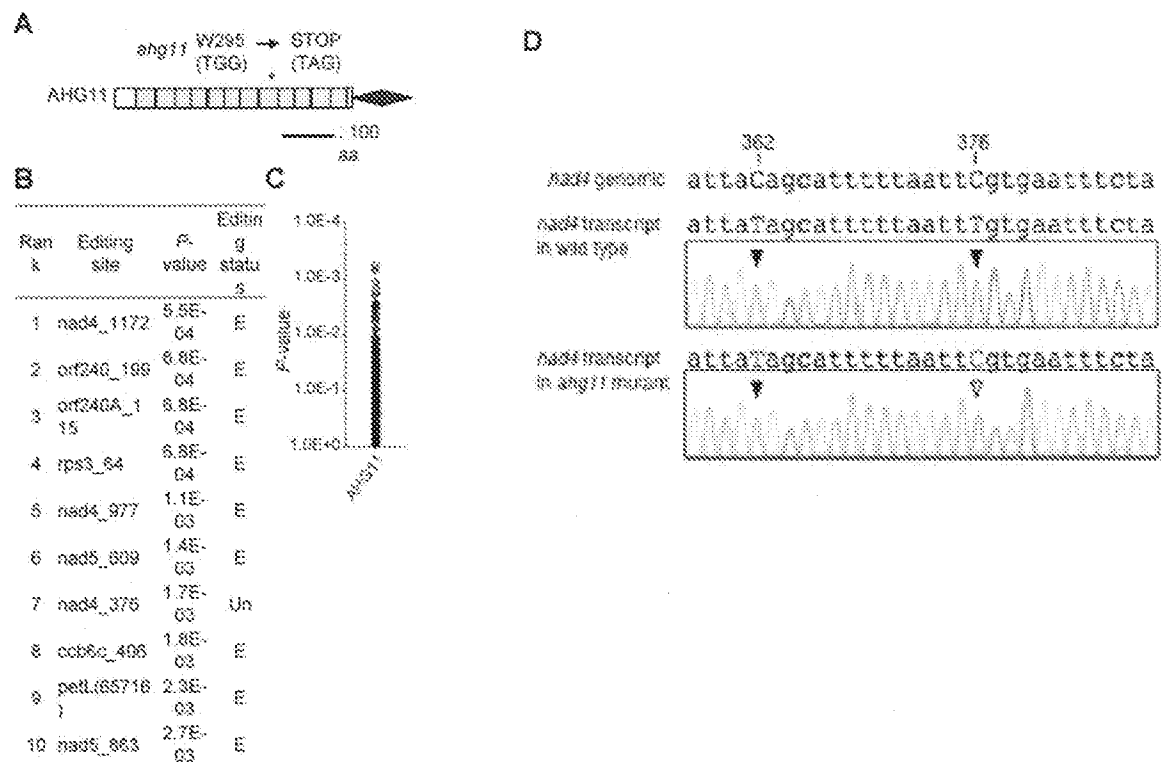
[図11]



[図12]



[図13]



[図14]

Protein	At code		At + Pp code*	
	P-value	Rank	P-value	Rank
CRR21	5.0e-8	1	1.0e-7	1
CRR4	6.0e-6	1	6.0e-6	1
LPA66	9.0e-5	21	4.0e-5	11
OTP80	>1.0e-5	ND	2.0e-5	8
OTP81	9.0e-5	74	5.0e-5	45
YS1	4.0e-5	21	3.0e-6	2

ND, not determined

[図15]

Table S3. List of scoring matrix for PPR code extracted from 214 PPR motifs in 24 *Arabidopsis* PPR proteins and 5 *Physcomitrella* PPR proteins

PPR code	occurrence of the code	Nucleotide occurrence frequency				Scoring matrix				Probability matrix (for MEME)				P (pt)	P (rv)	P (NS)
		A	G	U	C	A	G	U	C	A	G	U	C			
VND	21	0.80	0.80	1.17	5.33	0.24	0.33	13.80	8.83	0.02	0.02	0.87	0.28	2.2E-02	2.2E-02	5.9E-02
FND	12	1.83	1.00	8.83	2.33	-3.66	-2.33	-0.84	-1.87	0.15	0.08	0.57	0.19	7.7E-02	6.8E-02	1.2E-01
VTN	10	7.50	1.50	0.00	1.00	4.37	-0.52	-4.36	-1.46	0.75	0.15	0.00	0.10	1.2E-01	1.7E-01	1.1E-01
VNH	9	1.33	0.00	1.33	8.33	-1.28	-1.68	-2.33	4.28	0.15	0.00	0.15	0.70	2.5E-01	1.5E-01	2.2E-01
EGD	7	0.00	7.00	0.00	0.00	-2.55	5.49	-3.25	-1.84	0.00	1.00	0.00	0.00	1.7E-01	8.7E-01	8.9E-01
NNN	7	0.50	0.00	2.33	4.17	-1.39	-1.12	-0.25	2.73	0.07	0.00	0.33	0.60	1.0E-01	2.2E-01	6.1E-01
YTD	8	0.00	5.00	1.00	0.00	-1.83	3.82	-1.56	-1.43	0.00	0.83	0.17	0.00	8.5E-02	1.0E-01	1.0E-01
RTD	8	1.00	5.00	0.00	0.00	0.67	3.99	-2.12	-1.23	0.17	0.83	0.00	0.00	3.4E-02	1.3E-01	1.0E-01
FSN	8	4.67	0.33	0.33	0.67	3.10	-0.58	-1.86	-0.58	0.78	0.06	0.06	0.11	2.0E-02	1.0E-01	1.0E-01
ITD	8	0.00	3.00	1.00	1.00	1.57	1.89	-1.12	0.23	0.00	0.80	0.20	0.20	2.8E-01	6.5E-01	1.8E-01
YNG	8	0.00	0.80	3.00	1.50	-1.31	-0.84	1.17	0.48	0.00	0.10	0.80	0.30	2.4E-01	7.4E-02	6.8E-01
VSN	8	3.67	0.00	1.00	0.33	2.86	-0.84	-0.83	-0.50	0.73	0.00	0.20	0.07	5.4E-02	3.0E-01	5.8E-02
VNS	8	0.00	0.00	0.67	4.33	-1.31	-0.84	-1.16	3.31	0.00	0.00	0.13	0.87	4.7E-02	2.7E-01	1.0E-01
FTN	4	4.00	0.00	0.00	0.00	2.80	-0.84	-1.23	-1.00	1.00	0.00	0.00	0.00	7.4E-01	4.8E-01	4.3E-01
VSD	4	1.50	1.00	0.50	1.00	0.46	0.33	-0.96	0.18	0.08	0.25	0.13	0.25	9.2E-01	6.2E-01	1.0E-01
IND	4	1.00	0.00	3.00	0.00	-0.04	-0.87	1.54	-0.82	0.25	0.00	0.75	0.00	1.1E-01	3.2E-01	4.6E-01
ITN	4	4.00	0.00	0.00	0.00	2.86	-0.87	-1.46	-0.82	1.00	0.00	0.00	0.00	7.3E-01	4.9E-01	4.6E-01
IRD	4	0.00	0.00	3.00	1.00	-1.04	-0.67	1.34	0.18	0.00	0.00	0.75	0.25	1.1E-01	4.6E-01	3.2E-01
ISD	4	0.33	0.50	2.50	0.67	-0.71	-0.17	1.04	-0.15	0.08	0.12	0.63	0.17	3.8E-01	2.4E-01	4.6E-01
FPD	3	0.00	0.00	3.00	0.00	-1.04	-0.67	1.54	-0.82	0.00	0.00	1.00	0.00	8.3E-02	8.3E-02	8.3E-02
HNW	3	0.00	0.00	2.00	1.00	-0.78	-0.50	0.90	0.39	0.00	0.00	0.67	0.33	3.0E-01	8.9E-02	5.8E-01
VHW	3	0.50	0.00	1.00	1.50	-0.28	-0.50	-0.10	0.89	0.17	0.00	0.83	0.50	6.4E-01	2.5E-01	1.0E-01
VVE	3	1.00	0.00	1.50	0.50	0.22	-0.50	0.40	-0.11	0.33	0.00	0.50	0.17	6.4E-01	5.8E-01	2.5E-01
YPD	3	0.00	0.00	3.00	0.00	-0.78	-0.50	1.00	-0.51	0.00	0.00	1.00	0.00	2.0E-01	8.9E-02	8.9E-02
YTN	3	0.00	1.00	2.00	0.00	-0.78	0.50	0.90	-0.51	0.00	0.33	0.67	0.00	3.0E-01	5.8E-01	5.8E-01
EGN	3	2.00	0.00	1.00	0.00	1.22	-0.50	-0.10	-0.51	0.67	0.00	0.33	0.00	3.0E-01	5.8E-01	5.8E-01
LTN	3	0.00	3.00	0.00	0.00	-0.78	2.50	-1.10	-0.51	0.00	1.00	0.00	0.00	2.0E-01	8.9E-02	8.9E-02
END	3	0.00	1.00	1.50	0.50	-0.78	0.50	0.90	-0.51	0.00	0.33	0.50	0.17	6.4E-01	5.8E-01	1.0E-01
HRD	3	0.50	0.50	1.50	0.50	-0.28	0.00	0.40	-0.11	0.17	0.17	0.50	0.17	6.0E-01	5.6E-01	5.6E-01
IVD	3	1.00	0.00	0.00	2.00	0.22	-0.50	-1.10	1.39	0.33	0.00	0.00	0.67	3.0E-01	5.8E-01	5.8E-01
SNR	3	0.00	0.00	0.50	2.50	-0.78	-0.50	-0.50	1.89	0.00	0.00	0.17	0.83	1.5E-01	8.3E-02	2.5E-01
VVD	3	1.00	0.00	1.00	1.00	0.22	-0.50	-0.10	0.89	0.33	0.00	0.33	0.33	8.0E-01	5.6E-01	5.6E-01
FNW	3	0.33	1.00	0.67	1.00	-0.46	0.50	-0.43	0.39	0.11	0.33	0.22	0.33	9.4E-01	5.6E-01	5.6E-01
FTN	3	1.67	0.33	1.00	0.00	0.68	-0.17	-0.10	-0.51	0.55	0.11	0.33	0.00	5.0E-01	5.8E-01	1.8E-01
FTN	3	1.67	1.33	0.00	0.00	0.68	0.83	-1.10	-0.51	0.55	0.44	0.00	0.00	3.8E-01	8.3E-02	6.5E-01
ND	57	5.33	8.00	40.28	11.67	-3.55	-3.22	19.18	-9.32	0.09	0.11	0.79	0.20	2.2E-01	6.0E-01	1.3E-01
NN	27	3.83	3.50	7.16	15.49	-3.22	-1.05	-2.71	8.06	0.14	0.13	0.37	0.37	2.7E-01	2.7E-01	1.3E-01
TN	25	20.82	5.83	2.67	1.67	14.31	-0.37	-8.47	-3.46	0.83	0.15	0.11	0.07	8.9E-01	2.4E-01	2.5E-01
TD	21	3.28	19.67	2.67	2.33	-2.15	18.13	-5.91	-1.97	0.16	0.34	0.13	0.11	2.9E-01	2.9E-01	1.4E-01
SD	20	4.23	6.50	5.00	3.67	-0.39	3.13	-2.31	-0.43	0.04	0.33	0.25	0.18	8.5E-01	5.5E-01	5.4E-01
SN	15	10.33	0.33	3.33	1.00	8.42	-2.19	-2.15	-0.08	0.69	0.02	0.29	0.07	8.0E-01	1.0E-01	1.3E-01
GD	13	0.50	9.00	3.50	0.00	-2.80	6.81	-1.25	-2.86	0.04	0.68	0.27	0.00	1.7E-01	9.6E-02	1.7E-01
NS	13	0.83	0.50	4.67	7.00	2.58	-1.88	-0.09	4.83	0.06	0.04	0.36	0.54	8.6E-01	4.2E-01	5.9E-01
VN	12	0.67	1.33	5.00	2.00	0.59	-0.88	0.81	-0.46	0.31	0.11	0.42	0.17	4.3E-01	5.6E-01	1.2E-01
NW	10	1.15	1.33	3.67	3.83	1.45	-0.35	0.91	1.78	0.12	0.12	0.87	0.38	4.7E-01	1.1E-01	8.2E-01
PD	9	0.00	0.67	7.83	0.50	-2.35	-0.85	4.54	-1.34	0.00	0.07	0.87	0.06	2.3E-01	1.1E-01	2.7E-01
NT	8	1.00	1.50	1.67	5.83	-1.35	-0.01	-1.82	3.09	0.11	0.17	0.19	0.65	7.9E-02	8.9E-02	1.1E-01
GN	8	5.00	0.00	3.00	0.00	2.91	-1.35	0.67	-1.54	0.63	0.00	0.38	0.00	2.0E-01	4.8E-01	8.9E-01
MD	7	1.00	1.50	5.50	1.00	-0.83	0.32	2.34	-0.40	0.14	0.21	0.79	0.14	5.7E-01	9.1E-02	9.1E-02
VD	6	2.00	1.00	2.00	3.00	0.43	-0.01	-0.19	1.77	0.33	0.17	0.33	0.00	5.7E-01	2.5E-01	4.1E-01
VO	6	1.00	0.00	1.00	3.00	-0.51	-0.84	-0.35	1.98	0.20	0.00	0.20	0.60	2.8E-01	1.8E-01	6.5E-01
AD	4	0.33	3.00	0.67	0.00	-0.71	2.33	-0.30	-0.32	0.08	0.75	0.17	0.00	1.4E-01	1.8E-01	3.2E-01
GS	4	0.00	0.33	2.67	1.00	-1.04	-0.34	1.20	0.18	0.00	0.08	0.87	0.25	2.4E-01	9.0E-02	5.1E-01
VE	4	1.00	0.00	2.50	0.50	-0.04	-0.67	1.04	-0.32	0.25	0.00	0.63	0.13	3.2E-01	3.2E-01	1.3E-01
AN	4	0.67	0.00	2.67	0.67	-0.38	-0.67	1.20	-0.15	0.17	0.00	0.87	0.17	2.6E-01	1.8E-01	1.8E-01
CD	3	1.00	1.33	0.33	0.33	0.22	0.83	-0.77	-0.28	0.33	0.44	0.11	0.11	8.0E-01	3.3E-01	8.5E-01
IN	3	0.50	0.50	1.50	1.00	-0.25	-0.50	0.40	0.59	0.17	0.30	0.50	0.33	8.4E-01	2.5E-01	5.8E-01
TW	3	1.50	0.50	1.00	0.00	0.72	0.00	-0.10	-0.51	0.50	0.17	0.33	0.00	8.4E-01	5.6E-01	2.5E-01
VS	3	1.00	0.00	1.00	1.00	0.22	-0.50	-0.10	0.89	0.33	0.00	0.33	0.33	8.0E-01	5.6E-01	5.8E-01
ID	3	1.00	0.00	1.50	3.50	0.22	-0.50	0.40	2.89	0.33	0.30	0.50	1.17	8.6E-01	3.0E-01	6.8E-02
TS	3	0.00	0.50	1.50	1.00	-0.78	0.00	0.40	0.39	0.60	0.17	0.50	0.30	6.4E-01	2.5E-01	1.0E-01
IS	3	0.33	0.83	2.83	0.00	-0.45	0.33	1.74	-0.51	0.11	0.28	0.94	0.00	7.9E-02	2.8E-01	1.4E-01
VT	3	1.00	0.00	3.00	0.00	0.22	-0.50	1.80	-0.51	0.33	0.00	1.00	0.00	4.0E-01	2.0E-01	1.7E-01
Background frequency		0.26	0.17	0.37	0.20	0.00	0.00	0.00	0.00	0.26	0.17	0.37	0.20			

* The value indicates a probability of nucleotide occurrence for all RNA sequences surrounding the PPR editing sites used to construct the PPR code.

* The scoring matrix is estimated by a subtraction of the background frequency from the nucleotide occurrence frequency of the PPR code.

* The probability matrix is obtained from the nucleotide occurrence frequency.

[illegible]

[illegible]

[illegible]

[illegible]

id	name	email	phone	address	city	state	zip
1	John Doe	john.doe@example.com	555-123-4567	123 Main St	New York	NY	10001
2	Jane Smith	jane.smith@example.com	555-234-5678	456 Elm St	Los Angeles	CA	90001
3	Bob Johnson	bob.johnson@example.com	555-345-6789	789 Oak St	Chicago	IL	60601
4	Alice Brown	alice.brown@example.com	555-456-7890	101 Pine St	San Francisco	CA	94101
5	Charlie Davis	charlie.davis@example.com	555-567-8901	202 Cedar St	Phoenix	AZ	85001
6	Diana Evans	diana.evans@example.com	555-678-9012	303 Birch St	Philadelphia	PA	19101
7	Frank Green	frank.green@example.com	555-789-0123	404 Maple St	San Antonio	TX	78101
8	Grace Hall	grace.hall@example.com	555-890-1234	505 Walnut St	San Diego	CA	92101
9	Henry King	henry.king@example.com	555-901-2345	606 Elm St	San Jose	CA	95101
10	Ivy Lee	ivy.lee@example.com	555-012-3456	707 Oak St	San Luis Obispo	CA	93101
11	Jack Miller	jack.miller@example.com	555-123-4567	808 Pine St	San Marcos	CA	92101
12	Karen Wilson	karen.wilson@example.com	555-234-5678	909 Cedar St	San Mateo	CA	94401
13	Leo White	leo.white@example.com	555-345-6789	1010 Birch St	San Rafael	CA	94901
14	Mia Black	mia.black@example.com	555-456-7890	1111 Maple St	San Ramon	CA	94501
15	Noah Gray	noah.gray@example.com	555-567-8901	1212 Walnut St	San Bruno	CA	94066
16	Olivia Green	olivia.green@example.com	555-678-9012	1313 Elm St	San Clemente	CA	92673
17	Peter Hall	peter.hall@example.com	555-789-0123	1414 Oak St	San Diego	CA	92101
18	Quinn King	quinn.king@example.com	555-890-1234	1515 Pine St	San Diego	CA	92101
19	Rachel Lee	rachel.lee@example.com	555-901-2345	1616 Cedar St	San Diego	CA	92101
20	Sam Miller	sam.miller@example.com	555-012-3456	1717 Birch St	San Diego	CA	92101
21	Tina Wilson	tina.wilson@example.com	555-123-4567	1818 Maple St	San Diego	CA	92101
22	Uma White	uma.white@example.com	555-234-5678	1919 Walnut St	San Diego	CA	92101
23	Victor Black	victor.black@example.com	555-345-6789	2020 Elm St	San Diego	CA	92101
24	Wendy Gray	wendy.gray@example.com	555-456-7890	2121 Oak St	San Diego	CA	92101
25	Xavier Green	xavier.green@example.com	555-567-8901	2222 Pine St	San Diego	CA	92101
26	Yara Hall	yara.hall@example.com	555-678-9012	2323 Cedar St	San Diego	CA	92101
27	Zoe King	zoe.king@example.com	555-789-0123	2424 Birch St	San Diego	CA	92101
28	Adam Lee	adam.lee@example.com	555-890-1234	2525 Maple St	San Diego	CA	92101
29	Bella Miller	bella.miller@example.com	555-901-2345	2626 Walnut St	San Diego	CA	92101
30	Carl Wilson	carl.wilson@example.com	555-012-3456	2727 Elm St	San Diego	CA	92101
31	Dora White	dora.white@example.com	555-123-4567	2828 Oak St	San Diego	CA	92101
32	Ethan Black	ethan.black@example.com	555-234-5678	2929 Pine St	San Diego	CA	92101
33	Fiona Gray	fiona.gray@example.com	555-345-6789	3030 Cedar St	San Diego	CA	92101
34	Gavin Green	gavin.green@example.com	555-456-7890	3131 Birch St	San Diego	CA	92101
35	Hannah Hall	hannah.hall@example.com	555-567-8901	3232 Maple St	San Diego	CA	92101
36	Ian King	ian.king@example.com	555-678-9012	3333 Walnut St	San Diego	CA	92101
37	Jessica Lee	jessica.lee@example.com	555-789-0123	3434 Elm St	San Diego	CA	92101
38	Kyle Miller	kyle.miller@example.com	555-890-1234	3535 Oak St	San Diego	CA	92101
39	Laura Wilson	laura.wilson@example.com	555-901-2345	3636 Pine St	San Diego	CA	92101
40	Mason White	mason.white@example.com	555-012-3456	3737 Cedar St	San Diego	CA	92101
41	Natalie Black	natalie.black@example.com	555-123-4567	3838 Birch St	San Diego	CA	92101
42	Oscar Gray	oscar.gray@example.com	555-234-5678	3939 Maple St	San Diego	CA	92101
43	Pamela Green	pamela.green@example.com	555-345-6789	4040 Walnut St	San Diego	CA	92101
44	Quinn Hall	quinn.hall@example.com	555-456-7890	4141 Elm St	San Diego	CA	92101
45	Rachel King	rachel.king@example.com	555-567-8901	4242 Oak St	San Diego	CA	92101
46	Sam Lee	sam.lee@example.com	555-678-9012	4343 Pine St	San Diego	CA	92101
47	Tina Miller	tina.miller@example.com	555-789-0123	4444 Cedar St	San Diego	CA	92101
48	Uma Wilson	uma.wilson@example.com	555-890-1234	4545 Birch St	San Diego	CA	92101
49	Victor White	victor.white@example.com	555-901-2345	4646 Maple St	San Diego	CA	92101
50	Wendy Black	wendy.black@example.com	555-012-3456	4747 Walnut St	San Diego	CA	92101
51	Xavier Gray	xavier.gray@example.com	555-123-4567	4848 Elm St	San Diego	CA	92101
52	Yara Green						

[illegible]

id	name	email	phone	age	sex	height	weight
1	John Doe	john.doe@example.com	1234567890	30	M	180	75
2	Jane Smith	jane.smith@example.com	9876543210	25	F	165	60
3	Bob Johnson	bob.johnson@example.com	5555555555	40	M	175	80
4	Alice Brown	alice.brown@example.com	1111111111	28	F	170	65
5	Charlie Davis	charlie.davis@example.com	2222222222	35	M	185	70
6	Diana Prince	diana.prince@example.com	3333333333	22	F	160	55
7	Frank Miller	frank.miller@example.com	4444444444	45	M	190	85
8	Grace Wilson	grace.wilson@example.com	5555555555	20	F	155	50
9	Henry Taylor	henry.taylor@example.com	6666666666	38	M	178	72
10	Ivy Green	ivy.green@example.com	7777777777	27	F	168	62
11	Jack Black	jack.black@example.com	8888888888	32	M	182	78
12	Karen White	karen.white@example.com	9999999999	24	F	162	58
13	Leo Green	leo.green@example.com	1010101010	37	M	188	74
14	Mia Black	mia.black@example.com	2020202020	21	F	158	52
15	Noah White	noah.white@example.com	3030303030	42	M	192	88
16	Olivia Brown	olivia.brown@example.com	4040404040	26	F	172	68
17	Peter Davis	peter.davis@example.com	5050505050	34	M	184	76
18	Quinn Prince	quinn.prince@example.com	6060606060	19	F	152	48
19	Rachel Miller	rachel.miller@example.com	7070707070	41	F	186	82
20	Sam Wilson	sam.wilson@example.com	8080808080	23	M	164	56
21	Tina Taylor	tina.taylor@example.com	9090909090	36	F	176	70
22	Uma Green	uma.green@example.com	0101010101	29	F	166	64
23	Victor Black	victor.black@example.com	1212121212	31	M	180	77
24	Wendy White	wendy.white@example.com	2323232323	20	F	156	51
25	Xavier Brown	xavier.brown@example.com	3434343434	43	M	194	90
26	Yara Davis	yara.davis@example.com	4545454545	25	F	170	66
27	Zoe Prince	zoe.prince@example.com	5656565656	18	F	150	46
28	Adam Miller	adam.miller@example.com	6767676767	44	M	196	92
29	Bella Wilson	bella.wilson@example.com	7878787878	22	F	160	54
30	Carl Taylor	carl.taylor@example.com	8989898989	39	M	182	79
31	Dora Green	dora.green@example.com	9090909090	27	F	168	62
32	Ethan Black	ethan.black@example.com	0101010101	33	M	184	76
33	Fiona White	fiona.white@example.com	1212121212	21	F	158	52
34	Gavin Brown	gavin.brown@example.com	2323232323	46	M	198	94
35	Hannah Davis	hannah.davis@example.com	3434343434	24	F	172	68
36	Ian Prince	ian.prince@example.com	4545454545	17	M	148	44
37	Jessica Miller	jessica.miller@example.com	5656565656	47	F	188	86
38	Kyle Wilson	kyle.wilson@example.com	6767676767	23	M	164	56
39	Liam Taylor	liam.taylor@example.com	7878787878	40	M	186	84
40	Mia Green	mia.green@example.com	8989898989	28	F	170	70
41	Nora Black	nora.black@example.com	9090909090	19	F	152	48
42	Oscar White	oscar.white@example.com	0101010101	48	M	200	96
43	Pamela Brown	pamela.brown@example.com	1212121212	26	F	174	72
44	Quinn Davis	quinn.davis@example.com	2323232323	20	F	154	49
45	Ryan Prince	ryan.prince@example.com	3434343434	49	M	202	98
46	Sarah Miller	sarah.miller@example.com	4545454545	25	F	176	74
47	Timothy Wilson	timothy.wilson@example.com	5656565656	24	M	166	58
48	Uma Taylor	uma.taylor@example.com	6767676767	41	F	188	82
49	Victor Green	victor.green@example.com	7878787878	29	M	178	72
50	Wendy Black	wendy.black@example.com	8989898989	22	F	160	60
51	Xavier White	xavier.white@example.com	9090909090	50	M	204	100
52	Yara Brown	yara.brown@example.com	0101010101	27	F	172	70
53	Zoe Davis	zoe.davis@example.com	1212121212	21	F	156	51
54	Adam Prince	adam.prince@example.com	2323232323	51	M	206	102
55	Bella Miller	bella.miller@example.com	3434343434	26	F	174	72
56	Carl Wilson	carl.wilson@example.com	4545454545	25	M	168	64
57	Dora Taylor	dora.taylor@example.com	5656565656	42	F	190	88
58	Ethan Green	ethan.green@example.com	6767676767				

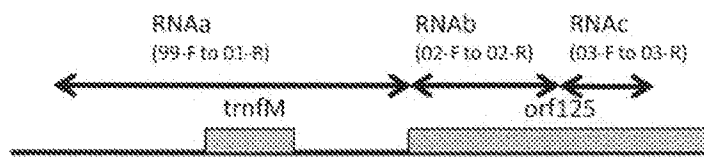
[illegible]

[illegible]

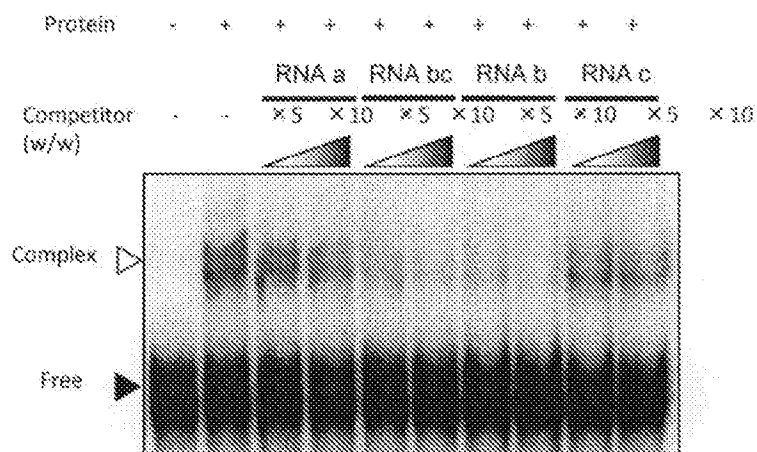
[illegible]

[図17]

A

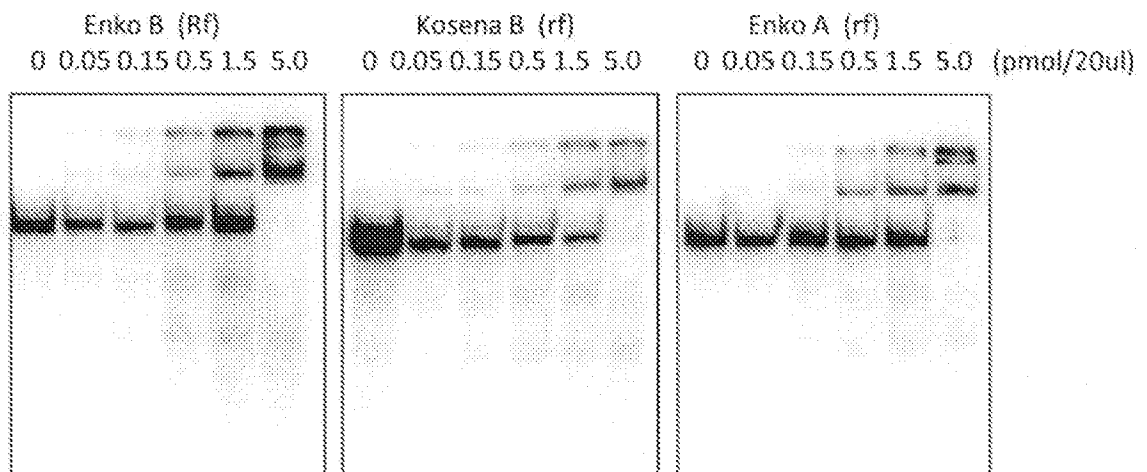


B

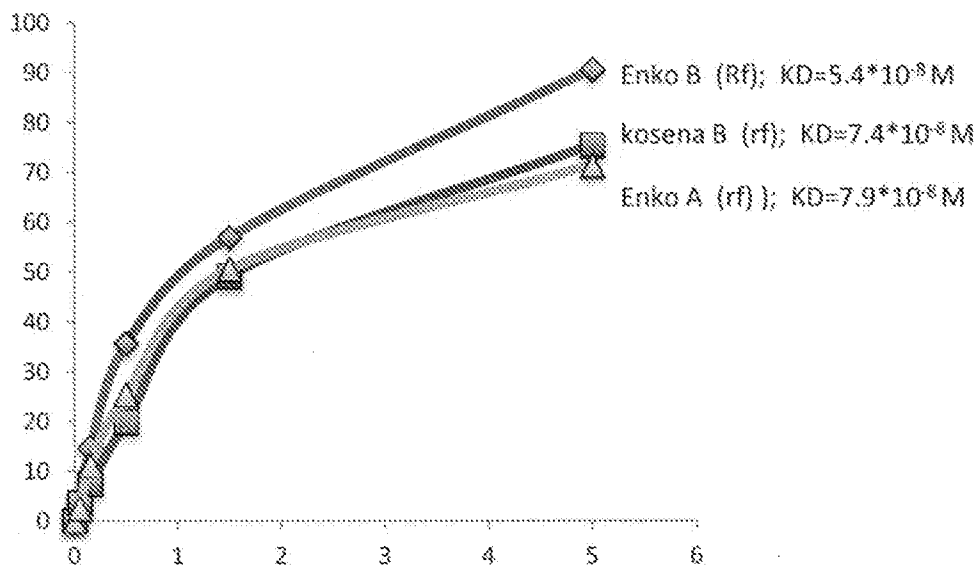


[図18]

A



B

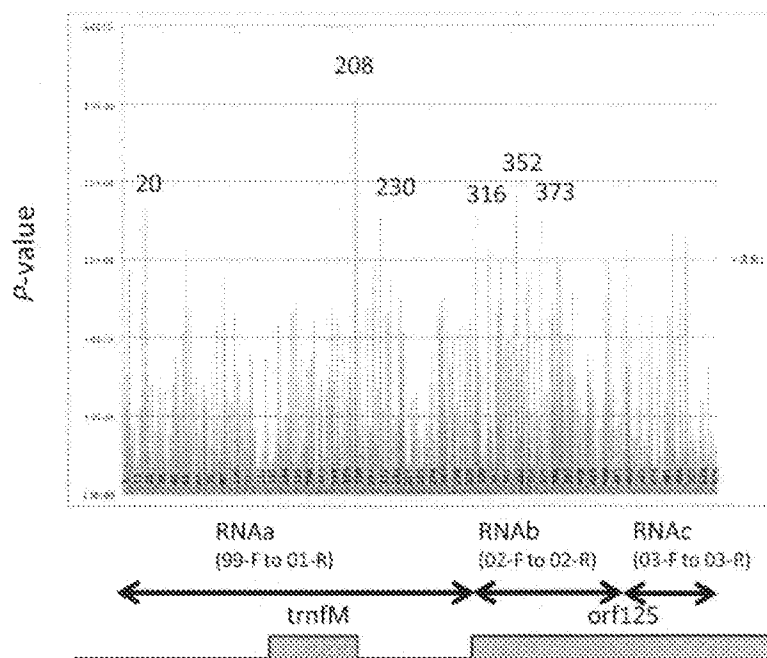


C

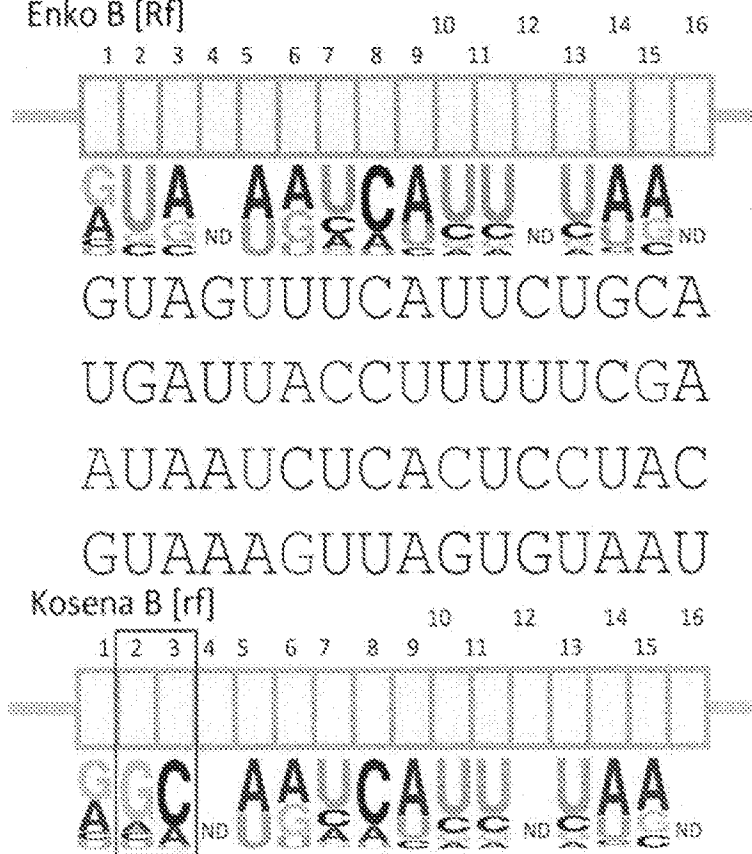
	EnkoB	kosenaB	EnkoA
208	8.399.E-06	4.620.E-04	6.299.E-02
316	2.702.E-04	2.667.E-04	2.630.E-04
352	1.372.E-04	7.168.E-03	3.535.E-01
373	2.711.E-04	5.013.E-03	4.007.E-01

A

large Rf (enko 8)

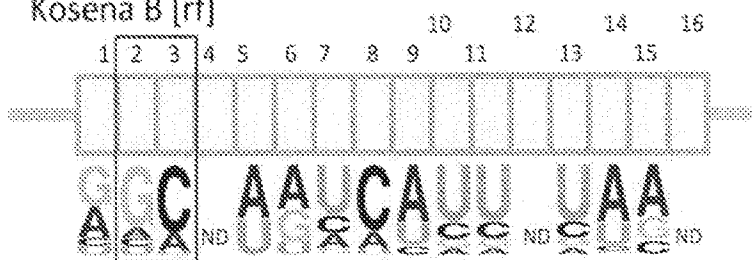


Enko B [Rf]

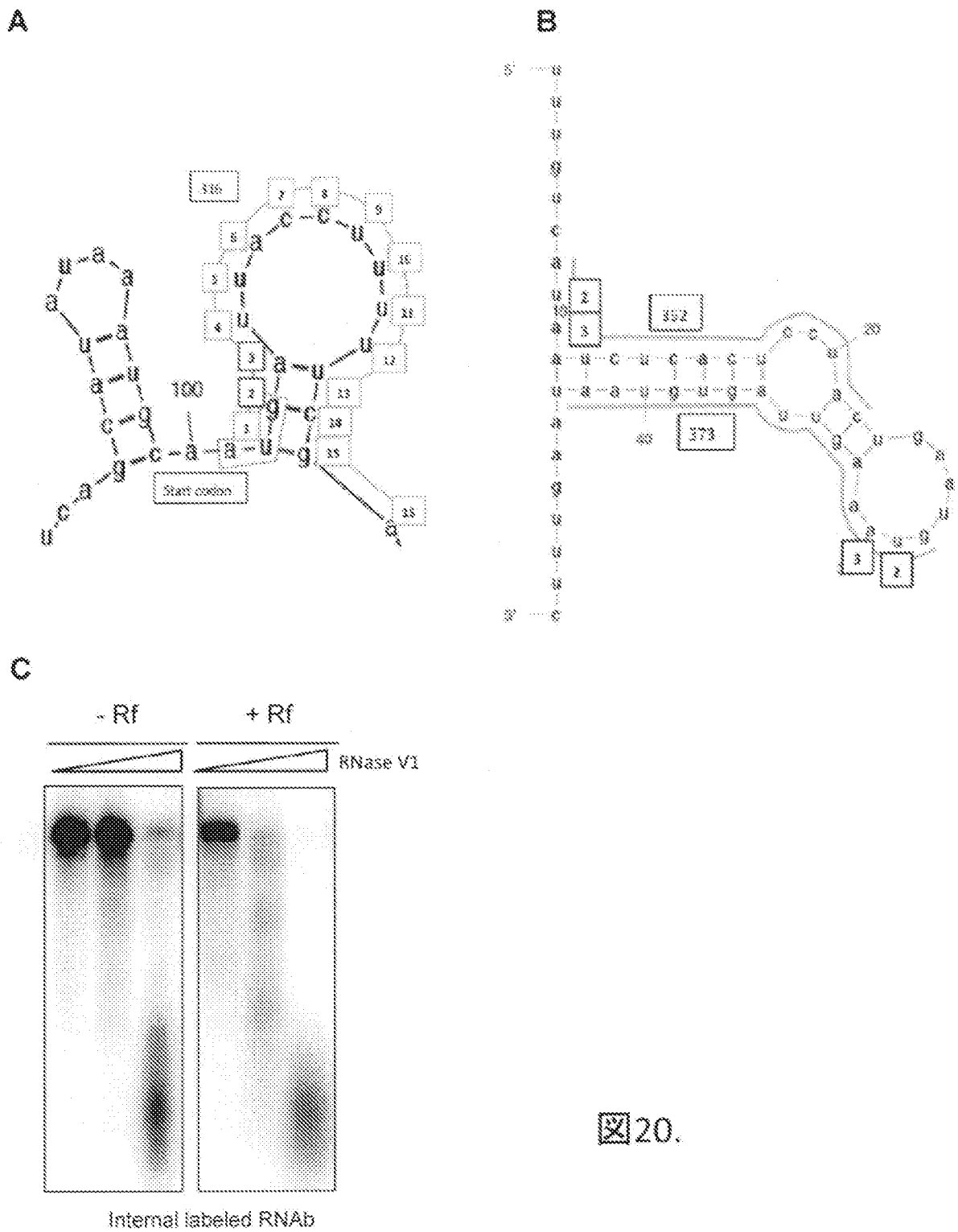


208: GUAGUUUCAUUCUGCA
316: UGAUUACCUUUUUCGA
352: AUAAUCUCACUCCUAC
373: GUAAAGUUAGUGUAAU

Kosena 8 (rf)



[図20]



[図20]

[図21-1]

Enko_B	MLAPVCFEESSESSPAASAARLPCTRSIRDTLAKAS--GESCAGFQGESLKLQSGPHKIK
Kosena_B	MLAPVCFEESSESSPAASAARLPCTRSIRDTLAKAS--GESCAGFQGESLKLQSGPHKIK
Comet_B	MLAPVCFEESSESSPAASAARLPCTRSIRDTLAKAS--GESCAGFQGESLKLQSGPHKIK
rrORF690-1	MLAPVCFEESSESSPAASAARLPCTRSIRDTLAKAS--GESCAGFQGESLKLQSGPHKIK
rrORF690-2	MLAPVCFEESSESSPAASAARLPCTRSIRDTLAKAS--GESCAGFQGESLKLQSGPHKIK
PC_PPR-A	MLAPVCFEESSESSPAASAARLPCTRSIRDTLAKAS--GESCAGFQGESLKLQSGPHKIK
Enko_A	MLAPVCFEESSESSPAASAARLPCTRSIRDTLAKAS--GESCAGFQGESLKLQSGPHKIK
Comet_A	MLAPVCFEESSESSPAASAARLPCTRSIRDTLAKAS--GESCAGFQGESLKLQSGPHKIK
Icicle_CA	MLAPVCFEESSESSPAASAARLPCTRSIRDTLAKAS--GESCAGFQGESLKLQSGPHKIK
PC_PPR-BL	MLAPVCFEESSESSPAASAARLPCTRSIRDTLAKAS--GESCAGFQGESLKLQSGPHKIK
PPR 1	
Enko_B	GLEDAIDLPSDMLSRPLPSVVDPCIKMGVVVMMERFDLVISLYQMERKQIRCDIYSFT
Kosena_B	GLEDAIDLPSDMLSRPLPSVVDPCIKMGVVVMMERFDLVISLYQMERKQIRCDIYSFT
Comet_B	GLEDAIDLPSDMLSRPLPSVVDPCIKMGVVVMMERFDLVISLYQMERKQIRCDIYSFT
rrORF690-1	GLEDAIDLPSDMLSRPLPSVVDPCIKMGVVVMMERFDLVISLYQMERKQIRCDIYSFT
rrORF690-2	GLEDAIDLPSDMLSRPLPSVVDPCIKMGVVVMMERFDLVISLYQMERKQIRCDIYSFT
PC_PPR-A	GLEDAIDLPSDMLSRPLPSVVDPCIKMGVVVMMERFDLVISLYQMERKQIRCDIYSFT
Enko_A	GLEDAIDLPSDMLSRPLPSVVDPCIKMGVVVMMERFDLVISLYQMERKQIRCDIYSFT
Comet_A	GLEDAIDLPSDMLSRPLPSVVDPCIKMGVVVMMERFDLVISLYQMERKQIRCDIYSFT
Icicle_CA	GLEDAIDLPSDMLSRPLPSVVDPCIKMGVVVMMERFDLVISLYQMERKQIRCDIYSFT
PC_PPR-BL	GLEDAIDLPSDMLSRPLPSVVDPCIKMGVVVMMERFDLVISLYQMERKQIRCDIYSFT
PPR 2	
Enko_B	ILINCFCSCKLPPALSTFGKITLGLHEDVTFPTLLHGLCDEVRVSEALDLPHQMPET
Kosena_B	ILINCFCSCKLPPALSTFGKITLGLHEDVTFPTLLHGLCDEVRVSEALDLPHQMPET
Comet_B	ILINCFCSCKLPPALSTFGKITLGLHEDVTFPTLLHGLCDEVRVSEALDLPHQMPET
rrORF690-1	ILINCFCSCKLPPALSTFGKITLGLHEDVTFPTLLHGLCDEVRVSEALDLPHQMPET
rrORF690-2	ILINCFCSCKLPPALSTFGKITLGLHEDVTFPTLLHGLCDEVRVSEALDLPHQMPET
PC_PPR-A	ILINCFCSCKLPPALSTFGKITLGLHEDVTFPTLLHGLCDEVRVSEALDLPHQMPET
Enko_A	ILINCFCSCKLPPALSTFGKITLGLHEDVTFPTLLHGLCDEVRVSEALDLPHQMPET
Comet_A	ILINCFCSCKLPPALSTFGKITLGLHEDVTFPTLLHGLCDEVRVSEALDLPHQMPET
Icicle_CA	ILINCFCSCKLPPALSTFGKITLGLHEDVTFPTLLHGLCDEVRVSEALDLPHQMPET
PC_PPR-BL	ILINCFCSCKLPPALSTFGKITLGLHEDVTFPTLLHGLCDEVRVSEALDLPHQMPET
PPR 3	
Enko_B	TCDPNNVTFTTLMHGLCREGRIVEAVALLDMMEDGLQPTQITYGTIVDMCKKMDTVSA
Kosena_B	TCDPNNVTFTTLMHGLCREGRIVEAVALLDMMEDGLQPTQITYGTIVDMCKKMDTVSA
Comet_B	TCDPNNVTFTTLMHGLCREGRIVEAVALLDMMEDGLQPTQITYGTIVDMCKKMDTVSA
rrORF690-1	TCDPNNVTFTTLMHGLCREGRIVEAVALLDMMEDGLQPTQITYGTIVDMCKKMDTVSA
rrORF690-2	TCDPNNVTFTTLMHGLCREGRIVEAVALLDMMEDGLQPTQITYGTIVDMCKKMDTVSA
PC_PPR-A	TCDPNNVTFTTLMHGLCREGRIVEAVALLDMMEDGLQPTQITYGTIVDMCKKMDTVSA
Enko_A	TCDPNNVTFTTLMHGLCREGRIVEAVALLDMMEDGLQPTQITYGTIVDMCKKMDTVSA
Comet_A	TCDPNNVTFTTLMHGLCREGRIVEAVALLDMMEDGLQPTQITYGTIVDMCKKMDTVSA
Icicle_CA	TCDPNNVTFTTLMHGLCREGRIVEAVALLDMMEDGLQPTQITYGTIVDMCKKMDTVSA
PC_PPR-BL	TCDPNNVTFTTLMHGLCREGRIVEAVALLDMMEDGLQPTQITYGTIVDMCKKMDTVSA
PPR 4	
Enko_B	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
Kosena_B	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
Comet_B	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
rrORF690-1	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
rrORF690-2	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
PC_PPR-A	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
Enko_A	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
Comet_A	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
Icicle_CA	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
PC_PPR-BL	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
PPR 5	
Enko_B	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
Kosena_B	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
Comet_B	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
rrORF690-1	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
rrORF690-2	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
PC_PPR-A	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
Enko_A	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
Comet_A	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
Icicle_CA	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
PC_PPR-BL	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
PPR 6	
Enko_B	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
Kosena_B	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
Comet_B	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
rrORF690-1	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
rrORF690-2	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
PC_PPR-A	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
Enko_A	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
Comet_A	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
Icicle_CA	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
PC_PPR-BL	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
PPR 7	
Enko_B	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
Kosena_B	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
Comet_B	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
rrORF690-1	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
rrORF690-2	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
PC_PPR-A	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
Enko_A	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
Comet_A	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
Icicle_CA	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
PC_PPR-BL	LNLLKXMEVSHIIPNNVTYSAILDGLCKDGRHSDAQNLFTEMQEKGIFFPIPTTNCMIV
PPR 8	
Enko_B	GFCSSGRWSEAAQALLQEMLEKXISPMVTYNALINAFVKEGKFFFAEELYDEMPLRGIIIF
Kosena_B	GFCSSGRWSEAAQALLQEMLEKXISPMVTYNALINAFVKEGKFFFAEELYDEMPLRGIIIF
Comet_B	GFCSSGRWSEAAQALLQEMLEKXISPMVTYNALINAFVKEGKFFFAEELYDEMPLRGIIIF
rrORF690-1	GFCSSGRWSEAAQALLQEMLEKXISPMVTYNALINAFVKEGKFFFAEELYDEMPLRGIIIF
rrORF690-2	GFCSSGRWSEAAQALLQEMLEKXISPMVTYNALINAFVKEGKFFFAEELYDEMPLRGIIIF
PC_PPR-A	GFCSSGRWSEAAQALLQEMLEKXISPMVTYNALINAFVKEGKFFFAEELYDEMPLRGIIIF
Enko_A	GFCSSGRWSEAAQALLQEMLEKXISPMVTYNALINAFVKEGKFFFAEELYDEMPLRGIIIF
Comet_A	GFCSSGRWSEAAQALLQEMLEKXISPMVTYNALINAFVKEGKFFFAEELYDEMPLRGIIIF
Icicle_CA	GFCSSGRWSEAAQALLQEMLEKXISPMVTYNALINAFVKEGKFFFAEELYDEMPLRGIIIF
PC_PPR-BL	GFCSSGRWSEAAQALLQEMLEKXISPMVTYNALINAFVKEGKFFFAEELYDEMPLRGIIIF

[図21-2]

	PPR 9	PPR 10
Enko_B	NITITSSMIDGFCQKRLDAAEHMFYLMATKGCSPNI	ITFTTLIDGYCGAKRIIDGMELL
Kosena_B	NITITSSMIDGFCQKRLDAAEHMFYLMATKGCSPNI	ITFTTLIDGYCGAKRIIDGMELL
Comet_B	NITITSSMIDGFCQKRLDAAEHMFYLMATKGCSPNI	ITFTTLIDGYCGAKRIIDGMELL
rrORF690-1	STVITYSSMIDGFCQKRLDAAEHMFYLMATKGCSPNI	ITFTTLIDGYCGAKRIIDGMELL
rrORF690-2	STVITYSSMIDGFCQKRLDAAEHMFYLMATKGCSPNI	ITFTTLIDGYCGAKRIIDGMELL
PC_FPR-A	NITITSSMIDGFCQKRLDAAEHMFYLMATKGCSPNI	ITFTTLIDGYCGAKRIIDGMELL
Enko_A	NITITSSMIDGFCQKRLDAAEHMFYLMATKGCSPNI	ITFTTLIDGYCGAKRIIDGMELL
Comet_A	NITITSSMIDGFCQKRLDAAEHMFYLMATKGCSPNI	ITFTTLIDGYCGAKRIIDGMELL
Icicle_CA	NITITSSMIDGFCQKRLDAAEHMFYLMATKGCSPNI	ITFTTLIDGYCGAKRIIDGMELL
PC_FPR-BL	NITITSSMIDGFCQKRLDAAEHMFYLMATKGCSPNI	ITFTTLIDGYCGAKRIIDGMELL
	PPR 11	PPR 12
Enko_B	HENTETGLVADTTTNTLIHQFCLVGLDNLAAALDLQSMISSGVCPE	IVTCNTLLDGLCIN
Kosena_B	HENTETGLVADTTTNTLIHQFCLVGLDNLAAALDLQSMISSGVCPE	IVTCNTLLDGLCIN
Comet_B	HENTETGLVADTTTNTLIHQFCLVGLDNLAAALDLQSMISSGVCPE	IVTCNTLLDGLCIN
rrORF690-1	HENTETGLVADTTTNTLIHQFCLVGLDNLAAALDLQSMISSGVCPE	IVTCNTLLDGLCIN
rrORF690-2	HENTETGLVADTTTNTLIHQFCLVGLDNLAAALDLQSMISSGVCPE	IVTCNTLLDGLCIN
PC_FPR-A	HENTETGLVADTTTNTLIHQFCLVGLDNLAAALDLQSMISSGVCPE	IVTCNTLLDGLCIN
Enko_A	HENTETGLVADTTTNTLIHQFCLVGLDNLAAALDLQSMISSGVCPE	IVTCNTLLDGLCIN
Comet_A	HENTETGLVADTTTNTLIHQFCLVGLDNLAAALDLQSMISSGVCPE	IVTCNTLLDGLCIN
Icicle_CA	HENTETGLVADTTTNTLIHQFCLVGLDNLAAALDLQSMISSGVCPE	IVTCNTLLDGLCIN
PC_FPR-BL	HENTETGLVADTTTNTLIHQFCLVGLDNLAAALDLQSMISSGVCPE	IVTCNTLLDGLCIN
	PPR 13	
Enko_B	GKIKDALEMPKAMQKSKMDLASHHFFNGVEFDM	LTYNILISGLINEGKFLAEELYSMP
Kosena_B	GKIKDALEMPKAMQKSKMDLASHHFFNGVEFDM	LTYNILISGLINEGKFLAEELYSMP
Comet_B	GKIKDALEMPKAMQKSKMDLASHHFFNGVEFDM	LTYNILISGLINEGKFLAEELYSMP
rrORF690-1	GKIKDALEMPKAMQKSKMDLASHHFFNGVEFDM	LTYNILISGLINEGKFLAEELYSMP
rrORF690-2	GKIKDALEMPKAMQKSKMDLASHHFFNGVEFDM	LTYNILISGLINEGKFLAEELYSMP
PC_FPR-A	GKIKDALEMPKAMQKSKMDLASHHFFNGVEFDM	LTYNILISGLINEGKFLAEELYSMP
Enko_A	GKIKDALEMPKAMQKSKMDLASHHFFNGVEFDM	LTYNILISGLINEGKFLAEELYSMP
Comet_A	GKIKDALEMPKAMQKSKMDLASHHFFNGVEFDM	LTYNILISGLINEGKFLAEELYSMP
Icicle_CA	GKIKDALEMPKAMQKSKMDLASHHFFNGVEFDM	LTYNILISGLINEGKFLAEELYSMP
PC_FPR-BL	GKIKDALEMPKAMQKSKMDLASHHFFNGVEFDM	LTYNILISGLINEGKFLAEELYSMP
	PPR 14	PPR 15
Enko_B	HRSIVPDTITYSMIDGLCKQSRIDEATQMFDSMKSSEFSPNI	VTFTTLINGYCKAGRV
Kosena_B	HRSIVPDTITYSMIDGLCKQSRIDEATQMFDSMKSSEFSPNI	VTFTTLINGYCKAGRV
Comet_B	HRSIVPDTITYSMIDGLCKQSRIDEATQMFDSMKSSEFSPNI	VTFTTLINGYCKAGRV
rrORF690-1	HRSIVPDTITYSMIDGLCKQSRIDEATQMFDSMKSSEFSPNI	VTFTTLINGYCKAGRV
rrORF690-2	HRSIVPDTITYSMIDGLCKQSRIDEATQMFDSMKSSEFSPNI	VTFTTLINGYCKAGRV
PC_FPR-A	HRSIVPDTITYSMIDGLCKQSRIDEATQMFDSMKSSEFSPNI	VTFTTLINGYCKAGRV
Enko_A	HRSIVPDTITYSMIDGLCKQSRIDEATQMFDSMKSSEFSPNI	VTFTTLINGYCKAGRV
Comet_A	HRSIVPDTITYSMIDGLCKQSRIDEATQMFDSMKSSEFSPNI	VTFTTLINGYCKAGRV
Icicle_CA	HRSIVPDTITYSMIDGLCKQSRIDEATQMFDSMKSSEFSPNI	VTFTTLINGYCKAGRV
PC_FPR-BL	HRSIVPDTITYSMIDGLCKQSRIDEATQMFDSMKSSEFSPNI	VTFTTLINGYCKAGRV
	PPR 16	
Enko_B	DGLELFCMGRGIVANNITITLIHQFCKVGNINGALDI	PQEMISSGVYFDTITIRNML
Kosena_B	DGLELFCMGRGIVANNITITLIHQFCKVGNINGALDI	PQEMISSGVYFDTITIRNML
Comet_B	DGLELFCMGRGIVANNITITLIHQFCKVGNINGALDI	PQEMISSGVYFDTITIRNML
rrORF690-1	DGLELFCMGRGIVANNITITLIHQFCKVGNINGALDI	PQEMISSGVYFDTITIRNML
rrORF690-2	DGLELFCMGRGIVANNITITLIHQFCKVGNINGALDI	PQEMISSGVYFDTITIRNML
PC_FPR-A	DGLELFCMGRGIVANNITITLIHQFCKVGNINGALDI	PQEMISSGVYFDTITIRNML
Enko_A	DGLELFCMGRGIVANNITITLIHQFCKVGNINGALDI	PQEMISSGVYFDTITIRNML
Comet_A	DGLELFCMGRGIVANNITITLIHQFCKVGNINGALDI	PQEMISSGVYFDTITIRNML
Icicle_CA	DGLELFCMGRGIVANNITITLIHQFCKVGNINGALDI	PQEMISSGVYFDTITIRNML
PC_FPR-BL	DGLELFCMGRGIVANNITITLIHQFCKVGNINGALDI	PQEMISSGVYFDTITIRNML
Enko_B	TGLWSKEELKRAVAMLEKLQMSMDLSFGQ-	
Kosena_B	TGLWSKEELKRAVAMLEKLQMSMDLSFGQ-	
Comet_B	TGLWSKEELKRAVAMLEKLQMSMDLSFGQ-	
rrORF690-1	TGLWSKEELKRAVAMLEKLQMSMDLSFGQ-	
rrORF690-2	TGLWSKEELKRAVAMLEKLQMSMDLSFGQ-	
PC_FPR-A	TGLWSKEELKRAVAMLEKLQMSMDLSFGQ-	
Enko_A	TGLWSKEELKRAVAMLEKLQMSMDLSFGQ-	
Comet_A	TGLWSKEELKRAVAMLEKLQMSMDLSFGQ-	
Icicle_CA	TGLWSKEELKRAVAMLEKLQMSMDLSFGQ-	
PC_FPR-BL	TGLWSKEELKRAVAMLEKLQMSMDLSFGQ-	

[図22]

		PPR motif No.																Amino acid
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Enko B (Rf, ND)	1	V	Y	V	V	I	V	F	V	I	I	T	V	Q	I	V	I	
	4	C	N	T	T	G	S	N	N	S	N	N	D	N	S	T	I	
	8	D	D	N	T	N	D	D	N	N	D	D	D	D	N	N	D	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Kosena B (rf, 99.4%)	1	V	Y	V	V	I	V	F	V	I	I	T	V	Q	I	V	I	
	4	C	T	N	T	G	S	N	N	S	N	N	D	N	S	T	I	
	8	D	D	N	T	N	D	D	N	N	D	D	D	D	N	N	D	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Comet B (Rf, 98.0%)	1	V	Y	V	V	I	V	F	V	I	I	T	V	Q	I	V	I	
	4	C	N	T	T	G	S	N	N	S	N	N	D	N	S	T	I	
	8	D	D	N	T	N	D	D	N	N	D	D	D	D	N	N	D	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Enko A (rf, 85.7%)	1	I	Y	V	L	I	V	V	V	I	F	V	V	L	I	V	I	
	4	N	T	T	T	G	S	N	N	N	T	N	N	N	S	N	I	
	8	D	D	D	D	N	N	N	N	D	N	D	D	D	N	N	D	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Comet A (rf, 87.6%)	1	I	Y	V	V	I	V	V	V	I	F	T	V	L	I	V	I	
	4	N	T	T	T	G	S	N	N	N	N	N	N	N	S	N	I	
	8	D	D	N	D	N	N	N	N	D	D	D	D	D	N	N	D	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Icicle CA (rf, 85.8%)	1	I	Y	V	V	I	V	F	V	I	F	V	V	L	I	V	I	
	4	C	T	N	T	G	S	S	N	S	N	T	N	N	S	N	I	
	8	N	D	N	N	D	D	D	N	D	N	N	D	D	N	N	D	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
mORF690-1 (?; 87.2%)	1	I	Y	A	I	I	V	F	V	V	I	T	V	L	F	V	I	
	4	N	N	N	T	G	S	T	S	S	N	N	N	N	N	N	I	
	8	D	D	N	D	N	N	N	S	N	D	D	D	D	N	N	D	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
mORF690-2 (?; 86.6%)	1	I	Y	A	I	I	V	F	V	V	I	T	V	L	V	V	I	
	4	N	N	N	T	G	S	T	S	S	N	N	N	N	N	N	I	
	8	D	D	N	D	N	N	N	S	N	D	D	D	D	N	N	D	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
PC_PPR-A (?; 85.8%)	1	V	Y	V	L	I	V	V	V	I	F	V	V	L	I	V	I	
	4	F	T	T	T	G	S	N	N	N	T	N	N	N	S	N	I	
	8	D	D	N	D	N	D	N	N	D	N	D	D	D	N	D	D	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
PC_PPR-BL (?; 82.8%)	1	V	Y	V	I	I	V	F	V	I	I	I	V	Q	I	V	I	
	4	F	T	S	T	G	S	N	S	N	S	T	N	N	N	T	R	
	8	N	S	N	N	N	N	N	N	D	N	N	D	D	D	N	D	

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2012/077274

A. CLASSIFICATION OF SUBJECT MATTER

C12N15/09(2006.01)i, C12Q1/68(2006.01)i, C07K14/00(2006.01)n

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12N15/09, C12Q1/68, C07K14/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CA/BIOSIS/MEDLINE/WPIDS (STN), UniProt/GeneSeq, JSTPlus (JDreamII)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
<u>X</u> A	GILLMAN JD et al., The petunia restorer of fertility protein is part of a large mitochondrial complex that interacts with transcripts of the CMS-associated locus, Plant J., 2007, Vol.49, No.2, pp.217-227	<u>7</u> 1-5, 9-16
Y	WO 2011/111829 A1 (Kyushu University), 15 September 2011 (15.09.2011), (Family: none)	1-5, 7, 9-16
Y	FUJII S et al., Selection patterns on restorer-like genes reveal a conflict between nuclear and mitochondrial genomes throughout angiosperm evolution, Proc. Natl. Acad. Sci. USA, 2011.01.25, Vol.108, No.4, pp.1723-1728	1-5, 7, 9-16

☒ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
09 January, 2013 (09.01.13)Date of mailing of the international search report
22 January, 2013 (22.01.13)Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2012/077274

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
<u>Y</u> A	DONG S et al., Specific and modular binding code for cytosine recognition in Pumilio/FBF (PUF) RNA-binding domains, J. Biol. Chem., 2011.07.29, Vol.286, No.30, pp.26732-26742	<u>1-5, 7</u> 9-16
<u>Y</u> A	Keizo KUNO et al., "Kosena-gata Saiboshitsu Daikon no Nensei Kaifuku Inshi to shite Hataraku PPR Tanpakushitsu no RNA Ketsugono ni Kansuru Kaiseki", Dai 52 Kai Proceedings of the Annual Meeting of the Japanese Society of Plant Physiologists, 11 March 2011 (11.03.2011), page 332	<u>9-16</u> 1-5, 7
<u>Y</u> A	WO 2002/088179 A1 (Mitsubishi Chemical Corp.), 07 November 2002 (07.11.2002), & JP 2003-289879 A & US 2004/0117868 A1 & US 2004/0088749 A1 & US 2010/0199376 A1 & EP 1382612 A1	<u>9-16</u> 1-5, 7
<u>Y</u> A	JP 2002-355041 A (Mitsubishi Chemical Corp.), 10 December 2002 (10.12.2002), (Family: none)	<u>9-16</u> 1-5, 7
<u>Y</u> A	KOIZUKA N et al., Genetic characterization of a pentatricopeptide repeat protein gene, orf687, that restores fertility in the cytoplasmic male-sterile Kosena radish, Plant J., 2003, Vol.34, No.4, pp.407-415	<u>9-16</u> 1-5, 7
P, X	BARKAN A et al., A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins, PLoS Genet., 2012.08.16, Vol.8, No.8, e1002910, doi:10.1371/journal.pgen.1002910	1-5, 7, 9-16
P, X	KOBAYASHI K et al., Identification and characterization of the RNA binding surface of the pentatricopeptide repeat protein, Nucleic Acids Res., 2012.03, Vol.40, No.6, pp.2712-2723, Epub. 2011.11.29	1-5, 7, 9-16
P, X	NAKAMURA T et al., Mechanistic insight into pentatricopeptide repeat proteins as sequence-specific RNA-binding proteins for organellar RNAs in plants, Plant Cell Physiol., 2012.07, Vol.53, No.7, pp.1171-1179, Epub. 2012.05.09	1-5, 7, 9-16
P, X	Yusuke YAGI et al., "Pentatricopeptide repeat (PPR) no Hairetsu Tokuiteki RNA Ninshiki Kiko", Proceedings of the 76th Annual Meeting of the Botanical Society of Japan, 14 September 2012 (14.09.2012), page 233	1-5, 7, 9-16

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2012/077274

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2009/113249 A1 (Tohoku University), 17 September 2009 (17.09.2009), & US 2011/0060033 A1 & EP 2258165 A1	1-5, 7, 9-16

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2012/077274

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.: 6,8
because they relate to subject matter not required to be searched by this Authority, namely:
Claims 6 and 8 pertain to method for treatment of the human body by surgery and thus relate to a subject matter on which this International Searching Authority is not required to carry out an international search under the provisions of PCT Article 17(2)(a)(i) and [PCT Rule 39.1(iv)].
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
(See extra sheet.)

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☒ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2012/077274

Continuation of Box No.III of continuation of first sheet (2)

The technical feature common to the invention set forth in claim 1 and the invention set forth in claim 9 is a PPR protein.

However, the common technical feature as said above cannot be considered to be a special technical feature, since the common technical feature does not make a contribution over the prior art in the light of the contents disclosed in the following documents 1 and 2.

Further, there is no other same or corresponding special technical feature between these inventions.

This opinion may be also applied to claims 10-16.

Accordingly, the following two inventions are involved in claims.

(Invention 1) the inventions of claims 1-5 and 7

(Invention 2) the inventions of claims 9-16

Document 1: WO 2011/111829 A1 (Kyushu University), 15 September 2011 (15.09.2011)

Document 2: Keizo KUNO et al., "Kosena-gata Saiboshitsu Daikon no Nensei Kaifuku Inshi to shite Hataraku PPR Tanpakushitsu no RNA Ketsugono ni Kansuru Kaiseki", Dai 52 Kai Proceedings of the Annual Meeting of the Japanese Society of Plant Physiologists, 11 March 2011 (11.03.2011), page 332

A. 発明の属する分野の分類（国際特許分類（I P C））

Int.Cl. C12N15/09(2006.01)i, C12Q1/68(2006.01)i, C07K14/00(2006.01)n

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（I P C））

Int.Cl. C12N15/09, C12Q1/68, C07K14/00

最小限資料以外の資料で調査を行った分野に含まれるもの

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

CA/BIOSIS/MEDLINE/WPIDS(STN), UniProt/GeneSeq, JSTPlus(JDreamII)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
<u>X</u> A	GILLMAN JD et al., The petunia restorer of fertility protein is part of a large mitochondrial complex that interacts with transcripts of the CMS-associated locus, Plant J., 2007, Vol. 49, No. 2, pp. 217-227	<u>7</u> 1-5, 9-16
Y	W0 2011/111829 A1（国立大学法人九州大学）2011.09.15, （ファミリーなし）	1-5, 7, 9-16

☒ C 欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの

「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）

「O」口頭による開示、使用、展示等に言及する文献

「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」同一パテントファミリー文献

国際調査を完了した日

0 9 . 0 1 . 2 0 1 3

国際調査報告の発送日

2 2 . 0 1 . 2 0 1 3

国際調査機関の名称及びあて先

日本国特許庁（I S A / J P）

郵便番号 1 0 0 - 8 9 1 5

東京都千代田区霞が関三丁目4番3号

特許庁審査官（権限のある職員）

西村 亜希子

4 N

3 4 3 5

電話番号 0 3 - 3 5 8 1 - 1 1 0 1 内線 3 4 8 8

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
Y	FUJII S et al., Selection patterns on restorer-like genes reveal a conflict between nuclear and mitochondrial genomes throughout angiosperm evolution, Proc. Natl. Acad. Sci. USA, 2011.01.25, Vol.108, No.4, pp.1723-1728	1-5, 7, 9-16
<u>Y</u> A	DONG S et al., Specific and modular binding code for cytosine recognition in Pumilio/FBF (PUF) RNA-binding domains, J. Biol. Chem., 2011.07.29, Vol.286, No.30, pp.26732-26742	<u>1-5, 7</u> 9-16
<u>Y</u> A	久野恵三ほか, コセナ型細胞質ダイコンの稔性回復因子として働く PPR 蛋白質の RNA 結合能に関する解析, 第 52 回日本植物生理学会年会要旨集, 2011.03.11, p.332	<u>9-16</u> 1-5, 7
<u>Y</u> A	WO 2002/088179 A1 (三菱化学株式会社) 2002.11.07, & JP 2003-289879 A & US 2004/0117868 A1 & US 2004/0088749 A1 & US 2010/0199376 A1 & EP 1382612 A1	<u>9-16</u> 1-5, 7
<u>Y</u> A	JP 2002-355041 A (三菱化学株式会社) 2002.12.10, (ファミリーなし)	<u>9-16</u> 1-5, 7
<u>Y</u> A	KOIZUKA N et al., Genetic characterization of a pentatricopeptide repeat protein gene, orf687, that restores fertility in the cytoplasmic male-sterile Kosena radish, Plant J., 2003, Vol.34, No.4, pp.407-415	<u>9-16</u> 1-5, 7
P, X	BARKAN A et al., A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins, PLoS Genet., 2012.08.16, Vol.8, No.8, e1002910, doi:10.1371/journal.pgen.1002910	1-5, 7, 9-16
P, X	KOBAYASHI K et al., Identification and characterization of the RNA binding surface of the pentatricopeptide repeat protein, Nucleic Acids Res., 2012.03, Vol.40, No.6, pp.2712-2723, Epub. 2011.11.29	1-5, 7, 9-16
P, X	NAKAMURA T et al., Mechanistic insight into pentatricopeptide repeat proteins as sequence-specific RNA-binding proteins for organellar RNAs in plants, Plant Cell Physiol., 2012.07, Vol.53, No.7, pp.1171-1179, Epub. 2012.05.09	1-5, 7, 9-16

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
P, X	八木祐介ほか, Pentatricopeptide repeat (PPR) の配列特異的 RNA 認識機構, 日本植物学会第76回大会研究発表記録, 2012. 09. 14, p. 233	1-5, 7, 9-16
A	WO 2009/113249 A1 (国立大学法人東北大学) 2009. 09. 17, & US 2011/0060033 A1 & EP 2258165 A1	1-5, 7, 9-16

第Ⅱ欄 請求の範囲の一部の調査ができないときの意見（第1ページの2の続き）

法第8条第3項（PCT17条(2)(a)）の規定により、この国際調査報告は次の理由により請求の範囲の一部について作成しなかった。

1. ☒ 請求項 6,8 は、この国際調査機関が調査をすることを要しない対象に係るものである。
つまり、
請求項6,8は手術による人体の処置方法に関するものであって、PCT17条(2)(a)(i)及び[PCT規則39.1(iv)]の規定により、この国際調査機関が国際調査をすることを要しない対象に係るものである。
2. ☐ 請求項 _____ は、有意義な国際調査をすることができる程度まで所定の要件を満たしていない国際出願の部分に係るものである。つまり、
3. ☐ 請求項 _____ は、従属請求の範囲であってPCT規則6.4(a)の第2文及び第3文の規定に従って記載されていない。

第Ⅲ欄 発明の単一性が欠如しているときの意見（第1ページの3の続き）

次に述べるようにこの国際出願に二以上の発明があるところの国際調査機関は認めた。
（特別ページ参照）

1. ☒ 出願人が必要な追加調査手数料をすべて期間内に納付したので、この国際調査報告は、すべての調査可能な請求項について作成した。
2. ☐ 追加調査手数料を要求するまでもなく、すべての調査可能な請求項について調査することができたので、追加調査手数料の納付を求めなかった。
3. ☐ 出願人が必要な追加調査手数料を一部のみしか期間内に納付しなかったため、この国際調査報告は、手数料の納付のあった次の請求項のみについて作成した。
4. ☐ 出願人が必要な追加調査手数料を期間内に納付しなかったため、この国際調査報告は、請求の範囲の最初に記載されている発明に係る次の請求項について作成した。

追加調査手数料の異議の申立てに関する注意

- ☐ 追加調査手数料及び、該当する場合には、異議申立手数料の納付と共に、出願人から異議申立てがあった。
- ☐ 追加調査手数料の納付と共に出願人から異議申立てがあったが、異議申立手数料が納付命令書に示した期間内に支払われなかった。
- ☒ 追加調査手数料の納付はあったが、異議申立てはなかった。

第 III 欄別紙

請求項 1 に係る発明と、請求項 9 に係る発明は、PPR 蛋白質という共通の技術的特徴を有している。しかしながら、このような共通の技術的特徴は、下記文献 1, 2 の開示内容に照らして、先行技術に対する貢献をもたらすものではないから、当該技術的特徴は、特別な技術的特徴であるとはいえない。また、これらの発明の間には、ほかに同一の又は対応する特別な技術的特徴は存在しない。このことは、請求項 1 0 - 1 6 についても同様である。

そして、請求の範囲には以下に示す 2 の発明が含まれる。

(発明 1) 請求項 1 - 5, 7 に係る発明

(発明 2) 請求項 9 - 1 6 に係る発明

文献 1 WO 2011/111829 A1 (国立大学法人九州大学) 2011. 09. 15

文献 2 久野恵三ほか, コセナ型細胞質ダイコンの稔性回復因子として働く PPR 蛋白質の RNA 結合能に関する解析, 第 5 2 回日本植物生理学会年会要旨集, 2011. 03. 11, p. 332