



(12) **BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ**

(19) **Cộng hòa xã hội chủ nghĩa Việt Nam (VN)**
CỤC SỞ HỮU TRÍ TUỆ

(11) 
1-0022337

(51)⁷ **C12Q 1/68**

(13) **B**

(21)	1-2011-02815	(22)	17.04.2009
(86)	PCT/CN2009/000411	17.04.2009	(87) WO2010/118559
(45)	25.11.2019	380	(43) 30.01.2012
(73)	National Defense Medical Center (TW)		
	No. 161, Sec. 6, Minquan E. Rd., Neihu Dist. Taipei City 114, Taiwan		
(72)	LAI, Hung-Cheng (TW)		
(74)	Công ty TNHH Trường Xuân (AGELESS CO.,LTD.)		

(54) **PHƯƠNG PHÁP SÀNG LỌC UNG THƯ**

(57) Sáng chế đề cập đến phương pháp sàng lọc ung thư bao gồm các bước sau:

(1) chuẩn bị mẫu bệnh phẩm; (2) phát hiện tỷ lệ metyl hóa của trình tự CpG trong ít nhất một gen đích trong ADN hệ gen của mẫu bệnh phẩm, trong đó gen đích gồm có PTPRR, ZNF582, PDE8B, và DBC1; và (3) xác định có ung thư hoặc thay đổi bệnh lý ung thư hay không trong mẫu bệnh phẩm dựa vào trạng thái metyl hóa trong ít nhất một gen đích; trong đó phương pháp xét nghiệm để phát hiện trạng thái metyl hóa là phản ứng chuỗi polymeaza đặc hiệu metyl hóa (MSP), phản ứng chuỗi polymeaza đặc hiệu metyl hóa định lượng (QMSP), giải trình tự bisunfit (BS), vi dây, khối phổ, phân tích sắc ký lỏng hiệu năng cao biến tính (DHPLC); và tương tự.

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến phương pháp sàng lọc ung thư, chất đánh dấu sinh học để sàng lọc ung thư và cách sử dụng chúng. Cụ thể, sáng chế đề cập đến phương pháp sàng lọc ung thư sử dụng ADN đã được metyl hóa làm chất đánh dấu sinh học.

Tình trạng kỹ thuật của sáng chế

Bệnh ung thư cổ tử cung là một trong những nguyên nhân chính gây tử vong ở phụ nữ trên toàn thế giới cũng như ở Đài Loan. Dựa vào số liệu thống kê do Tổ chức y tế thế giới (World Health Organization - WHO) tiến hành năm 2002, ung thư cổ tử cung là căn bệnh lớn thứ hai gây tử vong cho phụ nữ trên toàn thế giới sau ung thư vú. Sàng lọc ung thư cổ tử cung đều đặn là cách tốt nhất để phòng ngừa bệnh ung thư cổ tử cung. Sàng lọc ung thư cổ tử cung thông thường bao gồm hai phương pháp: phương pháp xét nghiệm tế bào cổ tử cung (Pap smear) được sử dụng nhiều nhất, và xét nghiệm phát hiện virus gây u nhú ở người (xét nghiệm HPV). Phương pháp sàng lọc tế bào gồm bước lấy mẫu sinh thiết từ cổ tử cung, kiểm tra dưới kính hiển vi để xác định có sự thay đổi bệnh lý ung thư trong tế bào biểu mô đã được bóc tách hay không, để phát hiện ung thư cổ tử cung sớm. Ngoài ra, xét nghiệm HPV bao gồm bước kiểm tra xem có virus gây u nhú ở người (Human Papilloma Virus – HPV) trong mẫu bệnh phẩm hay không bằng cách sử dụng phản ứng chuỗi polymeaza (polymerase chain reaction - PCR) hoặc kỹ thuật lai bắt giữ (Hybrid Capture).

Tuy nhiên, phương pháp xét nghiệm tế bào cổ tử cung vẫn có nhiều đặc điểm không mong muốn. Một trong số đó là yêu cầu lấy mẫu bởi bác sĩ, và phân tích bởi chuyên viên xét nghiệm/nhà nghiên cứu bệnh học, điều đó dẫn đến chi phí nhân sự cao và gây khó khăn trong việc đẩy mạnh xét nghiệm ở nhiều nước đang phát triển. Ngoài ra, phương pháp xét nghiệm tế bào cổ tử cung có tỷ lệ âm tính giả cao, gây chậm trễ trong việc chẩn đoán và điều trị thích hợp trước khi có sự thay đổi bệnh lý ung thư. Còn đối với xét nghiệm HPV, mặc dù có độ nhạy cao, phương pháp này vẫn có xu hướng tạo ra tỷ lệ dương tính giả cao, điều đó không chỉ gây ra lo lắng vô ích ở các bệnh nhân mà còn gây lãng phí nhiều tài nguyên y tế trong các xét nghiệm tiếp theo cho các bệnh nhân có kết quả dương tính giả. Theo đó, một trong số các đề tài quan trọng trong việc thúc đẩy xét nghiệm bệnh ung thư cổ

từ cung là tăng độ chính xác và tính thuận tiện của phương pháp xét nghiệm bệnh ung thư cổ tử cung.

Các đột biến mất đoạn gen từ lâu đã được xem như là yếu tố quan trọng trong việc tạo khối u. Trong một thời gian dài, tác giả đã được làm quen với ý tưởng là khả năng mã hóa của hệ gen nằm trong sự sắp xếp bốn bazơ A, T, G, X. Thuyết hai lần tác động (two-hit theory) được đề xuất sớm vào những năm 1970 chỉ ra rằng các đột biến đồng phát hoặc các đột biến mất đoạn của một số gen ức chế khối u cùng loại có thể gây ra hoặc dẫn đến phát triển bệnh ung thư. Tuy nhiên, thông tin bổ sung ảnh hưởng đến kiểu hình có thể được lưu trong bazơ được cải biến 5-methylcytosine. 5-methylcytosine được phát hiện ở động vật có vú trong trình tự lặp đảo 5'-CpG-3'. Hầu hết các cặp dinucleotide CpG được methyl hóa trong các tế bào động vật có vú trừ một số vùng được gọi là “đảo CpG”. Các đảo CpG là các vùng trình tự giàu GC- và CpG- có kích thước xấp xỉ 1 kb, thường được nằm trong các gen lân cận nhau và thường được tìm thấy gần vùng gen khởi động của các gen được biểu hiện rộng. Sự methyl hóa cytosine xảy ra sau khi tổng hợp ADN, bằng sự chuyển enzym của nhóm methyl từ thể cho methyl S-adenosylmethionine tới vị trí cacbon-5 của cytosine. Phản ứng enzym được thực hiện bởi ADN methyltransferase (DNMT). DNMT1 là methyltransferase chính trong động vật có vú, và có thể đáp ứng lại sự phục hồi sau sao chép của các vị trí bán methyl hóa để methyl hóa hoàn toàn, được xem là methyl hóa duy trì, trong đó DNMT3A và DNMT3B được cho là liên quan chủ yếu đến việc methyl hóa các vị trí mới, quá trình được gọi là methyl hóa riêng biệt.

Sự không methyl hóa ở các dinucleotide CpG, tức là, giảm methyl hóa nói chung, là những bất thường biểu sinh thứ nhất được định rõ trong các tế bào ung thư. Tuy nhiên, trong vài năm gần đây, đã rõ ràng rằng sự methyl hóa quá mạnh vị trí đặc biệt, ví dụ ở một số gen ức chế khối u, có liên quan đến sự mất chức năng có thể tạo ra các ưu điểm chọn lọc trong quá trình sinh sản gây ung thư. Sự methyl hóa dày đặc của các đảo CpG ở các vùng gen khởi động có thể gây ra sự tái tạo mô hình chất nhiễm sắc thông qua các sửa đổi histon với sự bất hoạt gen sau đó. Do đó, bên cạnh các đột biến mất đoạn nhiễm sắc thể hoặc các đột biến gen, sự bất hoạt biểu sinh của các gen ức chế khối u bởi sự methyl hóa quá mạnh vùng gen khởi động thường thấy được trong bệnh ung thư ở người.

Gần đây, các nghiên cứu dịch tễ học đã thể hiện mối quan hệ giữa mức folat trong huyết tương, nguồn chính của nhóm methyl, với việc nhiễm và làm sạch HPV. Các hiện tượng đa hình di truyền của các enzym trong chuyển hóa của chu kỳ methyl cũng đã được phát hiện cùng với sự phát triển của các thương tổn trong biểu mô cổ tử cung. Khi khái niệm về biểu sinh học phát triển, các nghiên cứu tìm ra mối liên hệ giữa sự methyl hóa ADN và ung thư cổ tử cung cũng đang bùng nổ. Các nghiên cứu về sự methyl hóa ADN trong bệnh ung thư cổ tử cung đang ngày một nhiều, điều đó cho thấy tiềm năng của việc sử dụng methyl hóa làm các chất đánh dấu trong việc sàng lọc ung thư cổ tử cung. Với sự giao thoa tự nhiên giữa di truyền và môi trường, sự phổ biến của việc methyl hóa trong các gen ức chế khối u thay đổi trong các gen khác nhau và các quần thể khác nhau. Khái niệm về các kiểu hình methyl hóa với các tập tính bệnh khác nhau đã được đề xuất với nhiều tranh cãi. Kiểu hình methyl hóa của ung thư cổ tử cung và sự tương tác của nó với các kiểu gen HPV vẫn còn chưa rõ. Các gen nào được methyl hóa đặc biệt trong ung thư cổ tử cung và có bao nhiêu gen cần thiết để thu được ứng dụng lâm sàng sẽ vẫn là một vấn đề đang được quan tâm trong thời gian tới.

Tác giả đã nộp các đơn yêu cầu cấp patent ở Đài Loan (Công bố đơn yêu cầu cấp patent Đài Loan số 200831900), Trung Quốc (Đơn yêu cầu cấp patent Trung Quốc số 200810094659.2), Malaysia (UI20085354) và Mỹ (Công bố đơn yêu cầu cấp patent Mỹ số 20080311570) (Sau đây được gọi là các đơn tiền thân). So sánh với các đơn tiền thân này, tác giả đã phát hiện ra một số chất đánh dấu sinh học mới trong sàng lọc ung thư và phương pháp sàng lọc ung thư sử dụng chúng.

Bản chất kỹ thuật của sáng chế

Một mục đích của sáng chế là đề xuất phương pháp sàng lọc ung thư là sàng lọc ung thư cổ tử cung dòng thứ nhất.

Mục đích khác của sáng chế là đề xuất phương pháp sàng lọc ung thư, khác biệt ở chỗ phương pháp này có thể được sử dụng không chỉ trong sàng lọc ung thư cổ tử cung dòng thứ nhất mà còn trong sàng lọc ung thư dòng thứ hai để hỗ trợ cho xét nghiệm HPV hoặc kết quả sàng lọc tế bào chưa chắc chắn để thu được hiệu quả sàng lọc ung thư cổ tử cung chính xác hơn.

Mục đích khác nữa của sáng chế là đề xuất phương pháp sàng lọc ung thư, khác biệt ở chỗ phương pháp này có thể được sử dụng trong sàng lọc ung thư cổ tử cung cũng như sàng lọc các loại ung thư khác (ví dụ: ung thư buồng trứng, ung thư ruột kết và tương tự) để hỗ trợ chẩn đoán mẫu bệnh phẩm bất thường.

Mục đích khác nữa của sáng chế là đề xuất chất đánh dấu sinh học, hoặc kit chứa chất đánh dấu sinh học đó để sàng lọc nguy cơ mắc bệnh ung thư. Phương pháp sàng lọc ung thư có thể đạt được các mục đích nêu trên của sáng chế bao gồm bước phát hiện tỷ lệ metyl hóa của gen đích (chất đánh dấu sinh học) của mẫu bệnh phẩm và sử dụng kết quả xét nghiệm làm chỉ số sàng lọc để chỉ ra nguy cơ mắc ung thư, và phương pháp này bao gồm các bước sau:

bước 1: chuẩn bị mẫu bệnh phẩm phân lập được;

bước 2: phát hiện tỷ lệ metyl hóa của trình tự CpG ở ít nhất một gen đích trong ADN hệ gen của mẫu bệnh phẩm nói trên, trong đó gen đích bao gồm DBC1, PDE8B, PTPRR và ZNF582; và

bước 3: xác định xem có ung thư hoặc thay đổi bệnh lý tiền ung thư hay không trong mẫu bệnh phẩm dựa trên tỷ lệ metyl hóa trong gen đích, hoặc sử dụng tỷ lệ metyl hóa nói trên làm dấu ấn chẩn đoán;

trong đó mẫu bệnh phẩm bao gồm mẫu sinh thiết được nạo từ cổ tử cung, mô ung thư buồng trứng, dịch cổ tử cung, máu, nước tiểu, phân, nước bọt, tế bào màng nhầy ở miệng, và mô ung thư hậu phẫu;

trong đó các phương pháp xét nghiệm trạng thái metyl hóa của trình tự CpG trong gen đích bao gồm phản ứng chuỗi polymeaza đặc hiệu metyl hóa (methylation-specific polymerase chain reaction - MSP), phản ứng chuỗi polymeaza đặc hiệu metyl hóa định lượng (quantitative methylation-specific polymerase chain reaction - QMSP), giải trình tự bisunfit (bisulfite sequencing - BS), vi dây, khối phổ, phân tích sắc ký lỏng hiệu năng cao biến tính (denaturing high-performance liquid chromatography - DHPLC);

trong đó gen đích DBC1 có trình tự nucleotit được nêu trong SEQ ID No. 1;

trong đó gen đích PDE8B có trình tự nucleotit như được nêu trong SEQ ID No. 2;

trong đó gen đích PTPRR có trình tự nucleotit như được nêu trong SEQ ID No. 3; và

trong đó gen đích ZNF582 có trình tự nucleotit như được nêu trong SEQ ID No. 4.

Ngoài ra, chỉ số sàng lọc và phương pháp sàng lọc được mô tả ở trên có thể còn được sử dụng trong sàng lọc ung thư cổ tử cung và ung thư ruột kết.

Sáng chế còn đề xuất phương pháp sàng lọc ung thư buồng trứng, khác biệt ở chỗ phương pháp này xét nghiệm trạng thái metyl hóa của gen đích của mẫu bệnh phẩm, và sử dụng trạng thái metyl hóa này làm chỉ số sàng lọc để xác định nguy cơ mắc ung thư buồng trứng, và phương pháp này bao gồm các bước sau:

bước 1: chuẩn bị mẫu bệnh phẩm phân lập được;

bước 2: phát hiện tỷ lệ metyl hóa của trình tự CpG ở ít nhất một gen đích trong ADN hệ gen của mẫu bệnh phẩm nói trên, trong đó gen đích bao gồm DBC1, PTPRR và ZNF582; và

bước 3: xác định xem có ung thư buồng trứng hoặc thay đổi bệnh lý ung thư hay không dựa vào tỷ lệ metyl hóa trong gen đích, hoặc sử dụng tỷ lệ metyl hóa nói trên của gen đích làm dấu hiệu chẩn đoán;

trong đó mẫu bệnh phẩm nói trên bao gồm mô ung thư buồng trứng, dịch cổ tử cung, máu, nước tiểu, và mô ung thư hậu phẫu;

trong đó các phương pháp xét nghiệm đối với tỷ lệ metyl hóa của trình tự CpG trong gen đích nói trên bao gồm phản ứng chuỗi polymeaza đặc hiệu metyl hóa (methylation-specific polymerase chain reaction - MSP), phản ứng chuỗi polymeaza đặc hiệu metyl hóa định lượng (quantitative methylation-specific polymerase chain reaction - QMSP), giải trình tự bisunfit (bisulfite sequencing - BS), vi dây, khối phổ, phân tích sắc ký lỏng hiệu năng cao biến tính (denaturing high-performance liquid chromatography - DHPLC), và giải trình tự ADN dựa trên sự giải phóng của pyrophosphat (pyrosequencing),

trong đó gen đích DBC1 có trình tự nucleotit như được nêu trong SEQ ID No. 1;

trong đó gen đích PTPRR có trình tự nucleotit như được nêu trong SEQ ID No. 3; và

trong đó gen đích ZNF582 có trình tự nucleotit như được nêu trong SEQ ID No. 4.

Thuật ngữ “mẫu bệnh phẩm” ở đây là mẫu bệnh phẩm phân lập được, và mẫu bệnh phẩm phân lập được này bao gồm mẫu sinh thiết được nạo từ cổ tử cung, dịch cổ tử cung, máu, nước tiểu, phân, nước bọt, tế bào màng nhầy ở miệng, dịch dạ dày, mật, tế bào biểu mô

cổ tử cung, mô ung thư hậu phẫu, hoặc mẫu bệnh phẩm phù hợp khác. Phương pháp sàng lọc ung thư theo sáng chế hữu dụng để xét nghiệm trạng thái methyl hóa của các gen đích trong các mẫu bệnh phẩm phân lập được nói trên, và sử dụng trạng thái methyl hóa này của chất đánh dấu sinh học làm chỉ số sàng lọc của các loại ung thư khác nhau. Phương pháp sàng lọc ung thư và chỉ số sàng lọc của chất đánh dấu sinh học được đề xuất bởi sáng chế có thể được sử dụng bởi các nhà nghiên cứu để thực hiện xét nghiệm trong phòng thí nghiệm.

Trong đó, “nguy cơ có ung thư” ở đây là chỉ phần trăm khả năng mắc ung thư, nguy cơ bị bệnh ung thư trong tương lai, hoặc sự có mặt của tế bào ung thư hay không.

Trong đó khi trạng thái methyl hóa của gen đích (chất đánh dấu sinh học) cao hơn so với đối chứng thích hợp, nguy cơ có ung thư càng cao.

Theo phương pháp sàng lọc nói trên, sáng chế còn đề xuất chất đánh dấu sinh học để sàng lọc ung thư hoặc kit chứa chất đánh dấu sinh học đó để sàng lọc trạng thái methyl hóa của gen đích (chất đánh dấu sinh học) trong mẫu bệnh phẩm phân lập được để xác định nguy cơ mắc ung thư.

Các đặc điểm và ưu điểm này của sáng chế sẽ được hiểu đầy đủ và được đánh giá đúng từ phần mô tả chi tiết dưới đây cùng với các hình vẽ kèm theo.

Mô tả vắn tắt các hình vẽ

Fig.1A thể hiện kết quả thu được bằng cách thực hiện phân tích giải trình tự bisunfit (bisulfite sequencing - BS) trên các mẫu bệnh phẩm cổ tử cung khác nhau sử dụng gen đích PTPRR theo phương pháp sàng lọc ung thư theo sáng chế;

Fig.1B thể hiện kết quả thu được bằng cách thực hiện phân tích giải trình tự bisunfit (bisulfite sequencing - BS) trên các mẫu bệnh phẩm cổ tử cung khác nhau sử dụng gen đích ZNF582 theo phương pháp sàng lọc ung thư theo sáng chế;

Fig.1C thể hiện kết quả thu được bằng cách thực hiện phân tích giải trình tự bisunfit (bisulfite sequencing - BS) trên các mẫu bệnh phẩm cổ tử cung khác nhau sử dụng gen đích PDE8B theo phương pháp sàng lọc ung thư theo sáng chế; và

Fig.1D thể hiện kết quả thu được bằng cách thực hiện phân tích giải trình tự bisunfit (bisulfite sequencing - BS) trên các mẫu bệnh phẩm cổ tử cung khác nhau sử dụng gen đích DBC1 theo phương pháp sàng lọc ung thư theo sáng chế.

Mô tả chi tiết phương án ưu tiên

Sáng chế sẽ được mô tả chi tiết hơn với các ví dụ dưới đây, nhưng sáng chế không bị giới hạn ở các ví dụ đó.

Ví dụ 1: Vật liệu và phương pháp

1. Vật liệu

Vật liệu xét nghiệm bao gồm một loạt các mẫu bệnh phẩm cổ tử cung, bao gồm: ung thư biểu mô tế bào hình vảy (squamous cell carcinoma - SCC, n = 20), ung thư tuyến (adenocarcinoma - AC, n = 20), và mẫu bệnh phẩm cổ tử cung bình thường (n = 10). Tất cả các mẫu bệnh phẩm cổ tử cung, mẫu bệnh phẩm buồng trứng và mẫu bệnh phẩm ung thư ruột kết này được lấy từ bệnh viện Tri-service General Hospital, Đài Bắc. Mỗi mẫu bệnh phẩm phải trải qua quá trình chiết tách ADN hệ gen bằng các phương tiện của Kit QIAamp ADN (QIAGEN), và được sử dụng để so sánh và phân tích điều kiện metyl hóa ADN trong quá trình metyl hóa ADN trên bình diện bộ gen. Trước khi phân tích, chất lượng của ADN hệ gen phải được kiểm tra bằng thiết bị phân tích Bioanalyzer (Agilent). Trong ví dụ này, 10µg ADN được cắt đoạn phải trải qua xét nghiệm MeDIP (Methyl DNA IP).

2. Phân tích sự metyl hóa ADN bằng các xét nghiệm MeDIP và đảo CpG thuộc vùng gen khởi động

ADN hệ gen được cắt thành từng đoạn có kích thước nằm trong khoảng từ 300 đến 1.000 bp bởi Bioruptor™ UCS-200 (Diagenode). Kết lắng miễn dịch qua đêm ở nhiệt độ 4°C với 30 µl kháng thể đa dòng kháng 5'-metyl xytosin (Abcam) với thể tích cuối cùng là 100 µl dung dịch đệm IP (0,15% SDS, 1% Triton X-100, 150mM NaCl, 1 mM EDTA, 0,5 mM EGTA, 10 mM Tris và 0,1% BSA). Ủ hỗn hợp với 120 µl Protein G Sepharosa (Amersham) trong thời gian 2 giờ ở nhiệt độ 4°C và rửa hai lần với 1 ml nước muối nồng độ thấp, dung dịch đệm có nồng độ muối cao, liti clorua và đệm TE. Sau đó, protein G được xử lý hai lần bằng dung dịch đệm rửa giải (1% SDS, 0,1M NaHCO₃) trong thời gian 15 phút ở nhiệt độ trong phòng và thu lại ADN đã metyl hóa bằng cách chiết phenol-clorofom, sau đó kết tủa bằng etanol. ADN đã metyl hóa được làm giàu và ADN đưa vào được khuếch đại bằng Kit khuếch đại toàn bộ hệ gen (Whole Genome Amplification - Sigma). ADN đã được làm giàu và ADN tổng số lần lượt được đánh dấu là Cy5 và Cy3, và được lai đồng thời với

các dãy đảo CpG thuộc vùng gen khởi động (CpG Island-Plus-Promoter Arrays) được thiết kế và tổng hợp bởi NimbleGen Systems, Inc. Dãy này chứa 385.000 50-75 bp oligonucleotit được xếp mỗi 100 bp ngang qua 24.659 vùng gen khởi động HG18 RefSeq (từ 800 bp ngược dòng đến 200 bp xuôi dòng) và 28.226 đảo CpG, thể hiện theo cặp.

Mỗi đặc điểm trên dãy có một tỷ lệ log 2 được định tỷ lệ tương ứng. Tỷ lệ log2 được tiêu chuẩn hóa tập trung ở dữ liệu tỷ lệ gần bằng không. Từ dữ liệu tỷ lệ log2 được định tỷ lệ, cửa sổ có độ dài cố định (500 bp) được đặt quanh mỗi chất dò liên tục và thử nghiệm Kolmogorov-Smirnov (KS) một phía được sử dụng để xác định các chất dò được rút ra từ sự phân phối tích cực hơn đáng kể so với phần còn lại của dãy. Kết quả của mỗi chất dò là trị số P $-\log_{10}$ từ thử nghiệm KS tạo cửa sổ xung quanh chất dò đó (Scacheri et al, 2006). Nhóm tác giả phát hiện các đỉnh cực đại được làm giàu bằng cách dò tìm trong ít nhất hai chất dò trên ngưỡng trị số P tối thiểu ($-\log_{10}$) là 2. Các đỉnh cực đại trong vòng 300 bp của mỗi chất dò được nhập lại với nhau. Các vùng methyl hóa khác nhau giữa SCC/AC và cổ tử cung bình thường trong 2500 bp ở dòng cao hơn của các điểm sao chép được chọn để xác nhận trong tương lai. Cuối cùng, dữ liệu trị số p được theo dõi sử dụng SignalMap (NimbleGen).

3. Sửa đổi bisulfit, phản ứng chuỗi polymeaza đặc hiệu methyl hóa (MSP) và giải trình tự bisulfit (BS)

Kit sửa đổi ADN (Chemicon, Temecula, CA) được sử dụng hướng dẫn của nhà sản xuất để chuyển hóa 1 μ g phân lượng ADN hệ gen với natri bisulfit để bảo quản các xytosin đã methyl hóa. Phần kết lắng cuối cùng được rửa bằng 70 μ l đệm TE được làm ấm trước cho quá trình MSP.

ADN bình thường của máu ngoại vi ở người được đưa vào và trải qua sự sửa đổi bisulfit, được sử dụng làm nhóm đối chứng có trình tự gen khởi đầu chưa bị methyl hóa.

MSP được thực hiện theo kỹ thuật đã biết. Tóm lại, 1 μ l ADN đã sửa đổi được khuếch đại sử dụng các đoạn mồi MSP (Bảng 1) được xác nhận đặc hiệu về các trình tự gen được methyl hóa có mặt trong ADN được chuyển hóa bisulfit. MSP được thực hiện trong thể tích tổng cộng 25 μ l chứa 1 μ l ADN mẫu đã sửa đổi, 1,5 pmol mỗi đoạn mồi, 0,2 mmol/L deoxynucleotit triphosphat, và 1 đơn vị polymeraza Gold *Taq* ADN (Applied Biosystems,

Foster City, CA). Đầu tiên, các phản ứng MSP phải được ủ tại nhiệt độ 95°C trong thời gian 5 phút, tiếp theo là 35 chu kỳ ủ ở nhiệt độ 95°C trong thời gian 30 giây, và ủ ADN ở nhiệt độ thích hợp trong thời gian 30 giây và ở nhiệt độ 72°C trong thời gian 30 giây. Cuối cùng, thực hiện kéo dài ở nhiệt độ 72°C trong thời gian 5 phút. Các sản phẩm khuếch đại gen được nhìn thấy trên gel agarosa 2,5% chứa ethidi bromua và được chiếu tia UV.

Bảng 1. Trình tự của các đoạn mồi MSP

Gen	Đoạn mồi	Trình tự	
DBC1	M Xuôi (F')	5' gtttcgctgcttttcgttcggagatc 3'	(SEQ ID No. 5)
	M Ngược (R')	5' gctctcgctctcgctattactcgct 3'	(SEQ ID No. 6)
PDE8B	M Xuôi (F')	5' tgtgtatgcgcgttttcgttc 3'	(SEQ ID No. 7)
	M Ngược (R')	5' acctatatatccgccgctccgctc 3'	(SEQ ID No. 8)
PTPRR	M Xuôi (F')	5' cggcgttggtatgttttagtagtc 3'	(SEQ ID No. 9)
	M Ngược (R')	5' aattacgaataaaaaaacaaaaacgctc 3'	(SEQ ID No. 10)
ZNF582	M Xuôi (F')	5' tgacgggtttttgtttatcggttattc 3'	(SEQ ID No. 11)
	M Ngược (R')	5' cgaacgcaaacgtacctacgc 3'	(SEQ ID No. 12)

M: các đoạn mồi có thể nhận biết đặc hiệu các trình tự gen đã metyl hóa có trong ADN đã chuyển hóa bisulfit

Tất cả các dữ liệu MSP có nguồn gốc từ ít nhất hai sự sửa đổi ADN độc lập nhau. Sự không có tín hiệu trong hai thí nghiệm giống nhau được ghi lại là âm tính khi metyl hóa. Sản phẩm PCR thu được bằng cách sử dụng đoạn mồi MSP (M) có thể nhận biết đặc hiệu trình tự gen đã metyl hóa được đưa vào vector pCR4-TOPO (Invitrogen, Carlsbad, CA). 5 dòng độc lập được lựa chọn và trải qua quá trình giải trình tự bisulfit (bisulfite sequencing-BS). Các đoạn mồi được sử dụng trong giải trình tự bisulfit (BS) được liệt kê trong Bảng 2. Quá trình giải trình tự bisulfit được thực hiện sử dụng bộ giải trình tự tự động 377 (377 Auto-sequencer) (Applied Biosystems, Foster City, CA). Các kết quả giải trình tự được thể hiện trong danh mục trình tự gen. Các số trình tự trong giải trình tự bisulfit lần lượt là: DBC1_BS (SEQ ID No. 21), PDE8B_BS (SEQ ID No. 22), PTPRR_BS (SEQ ID No. 23) và ZNF582_BS (SEQ ID No. 24).

Bảng 2. Các trình tự của các đoạn môi giải trình tự bisulfit

Gen	Đoạn môi	Trình tự	
DBC1	Xuôi (F')	5'ggtaagtttttttyggygtagtt 3'	(SEQ ID No. 13)
	Ngược (R')	5'tactccctctacctcccractctctc 3'	(SEQ ID No. 14)
PDE8B	Xuôi (F')	5'tgtggygtagaggattattagtttgg 3'	(SEQ ID No. 15)
	Ngược (R')	5'ctaaaaacraaccatccctc 3'	(SEQ ID No. 16)
PTPRR	Xuôi (F')	5'ggaatttttttgaaattttttgtt 3'	(SEQ ID No. 17)
	Ngược (R')	5'ccccacttcaaataaaatactattaaaaaaac 3'	(SEQ ID No. 18)
ZNF582	Xuôi (F')	5'tagtgayggtttttgtttattyggttatt 3'	(SEQ ID No. 19)
	Ngược (R')	5'taaacrtaaaaacaacaaccracct 3'	(SEQ ID No. 20)

Ví dụ 2. Sàng lọc gen đích được metyl hóa trong ung thư cổ tử cung

Sau khi được sàng lọc bởi các dãy đảo CpG thuộc vùng gen khởi động, bốn gen đích có thể được metyl hóa cao trong tế bào ung thư cổ tử cung được sàng lọc, lần lượt cụ thể là, DBC1 (SEQ ID No. 1), PDE8B (SEQ ID No. 2), PTPRR (SEQ ID No. 3) và ZNF582 (SEQ ID No. 4). Thông tin chi tiết của chúng được thể hiện trong Bảng 3. Có thể thấy từ Bảng 3 là trong số bốn gen này, trừ DBC1 đã được biết đến gắn liền với bệnh ung thư bàng quang, có ít nghiên cứu cho đến nay bộc lộ mối quan hệ giữa các gen này với bệnh ung thư cổ tử cung. Bảng 3. Các đặc điểm của các gen được metyl hóa trong ung thư cổ tử cung được xác định sử dụng vi dãy đảo CpG thuộc vùng gen khởi động (CpG island-Plus-Promoter).

Gen	Mã số trong UniGene	Vị trí nhiễm sắc thể	Tên đầy đủ	SEQ ID số
DBC1	NM_014618	9q32-q33	Được loại bỏ trong ung thư bàng quan 1	SEQ ID No. 1
PDE8B	NM_003719	5q14.1	phosphodiesteraza 8B	SEQ ID No. 2

Gen	Mã số trong UniGene	Vị trí nhiễm sắc thể	Tên đầy đủ	SEQ ID số
PTPRR	NM_002849	12q15	protein tyrosin phosphatase, kiểu thụ thể, R	SEQ ID No. 3
ZNF582	NM_144690	19q13.43	Protein in đầu kẽm 582	SEQ ID No. 4

Ví dụ 3. Phân tích trạng thái methyl hóa của gen đích trong mẫu bệnh phẩm cổ tử cung bằng giải trình tự bisulfite (BS)

Gen đích: PTPRR

Các nhóm mẫu bệnh phẩm:

1. HeLa_0: dòng tế bào ung thư cổ tử cung HeLa;
2. HeLa_10D: dòng tế bào ung thư cổ tử cung HeLa (HeLa_0) được xử lý với chất ức chế ADN methyltransferaza 5'-aza-2'-deoxyxitydin 10 μ M (AZC);
3. HeLa_DT: dòng tế bào ung thư cổ tử cung HeLa (HeLa_0) được xử lý với chất ức chế ADN methyltransferaza 5'-aza-2'-deoxyxitydin 10 μ M (AZC) và TSA 0,33 μ M (Sigma Chemical Co., St. Louis, MO);
4. AC: mẫu bệnh phẩm ung thư tuyến cổ tử cung;
5. SCC: mẫu bệnh phẩm ung thư biểu mô tế bào cổ tử cung hình vảy;
6. Nhóm đối chứng (bình thường): ADN huyết thuộc cổ tử cung bình thường làm nhóm đối chứng không methyl hóa.

Mỗi mẫu bệnh phẩm nêu trên phải trải qua quá trình sửa đổi bisulfite, và sau đó là giải trình tự bisulfite để phân tích xem gen đích (PTPRR) trong mỗi mẫu bệnh phẩm đã bị methyl hóa cao hay không. Kết quả được thể hiện trên Fig.1, với phần màu đen là chỉ vùng methyl hóa và phần màu trắng là chỉ vùng chưa methyl hóa. Gen đích PTPRR không bị methyl hóa trong nhóm đối chứng và mẫu bệnh phẩm ung thư tuyến, trong khi gen đích PTPRR bị methyl hóa cao trong dòng tế bào ung thư cổ tử cung HeLa (HeLa_0) và mẫu bệnh phẩm ung thư biểu mô tế bào cổ tử cung hình vảy (SCC). Do đó, tỷ lệ methyl hóa của PTPRR có thể

được sử dụng làm chất đánh dấu sinh học để sàng lọc xem có ung thư cổ tử cung hay không.

Ngoài ra, để xác định có tỷ lệ methyl hóa của gen đích trong mẫu bệnh phẩm ung thư cổ tử cung có thể được điều qua trong quá trình methyl hóa ADN, dòng tế bào ung thư cổ tử cung HeLa được xử lý với 10 μ M chất ức chế ADN methyltransferaza 5'-aza-2'-deoxyxytidin (AZC) (Sigma Chemical Co.). Mẫu ADN được tách từ dòng tế bào này phải trải qua quá trình sửa đổi bisulfit và sau đó là giải trình tự bisulfit (BS). Kết quả được thể hiện trên Fig.1A chỉ ra rằng, so với mẫu bệnh phẩm ung thư biểu mô tế bào cổ tử cung hình vảy (SCC) và dòng tế bào ung thư cổ tử cung HeLa (HeLa_0), gen đích PTPRR trong dòng tế bào ung thư cổ tử cung HeLa đã được xử lý AZC (HeLa_10D) có một phần của vùng được giải methyl hóa.

Tricostatin A (TSA) là chất ức chế histon deacetylaza (HDAC), và có thể được sử dụng cho tỷ lệ methyl hóa thấp hơn. Dòng tế bào ung thư cổ tử cung HeLa (HeLa_DT) đã được xử lý với cả AZC và TSA. Kết quả được thể hiện trên Fig.1 chỉ ra rằng, so với mẫu bệnh phẩm ung thư biểu mô tế bào cổ tử cung hình vảy (SCC) và dòng tế bào ung thư cổ tử cung HeLa (HeLa_0), gen đích PTPRR trong dòng tế bào ung thư cổ tử cung HeLa được xử lý AZC- và TSA (HeLa_DT) đã được giải methyl hóa cao.

Tóm lại, các kết quả nêu trên đã xác nhận gen đích PTPRR trong mẫu bệnh phẩm ung thư cổ tử cung đã được methyl hóa thông qua sự methyl hóa ADN.

Gen đích: ZNF582

Các nhóm mẫu bệnh phẩm:

1. HeLa_0: dòng tế bào ung thư cổ tử cung HeLa;
2. HeLa_10D: dòng tế bào ung thư cổ tử cung HeLa (HeLa_0) được xử lý với chất ức chế ADN methyltransferaza 5'-aza-2'-deoxyxytidin 10 μ M (AZC);
3. HeLa_DT: dòng tế bào ung thư cổ tử cung HeLa (HeLa_0) được xử lý với cả chất ức chế ADN methyltransferaza 5'-aza-2'-deoxyxytidin 10 μ M (AZC) và TSA 0,33 μ M;
4. SCC 1: mẫu bệnh phẩm ung thư biểu mô tế bào cổ tử cung hình vảy 1;
5. SCC 2: mẫu bệnh phẩm ung thư biểu mô tế bào cổ tử cung hình vảy 2;
6. Nhóm đối chứng (bình thường): ADN huyết ở cổ tử cung bình thường làm nhóm đối chứng chưa methyl hóa;

Trong đó mẫu bệnh phẩm ung thư biểu mô tế bào cổ tử cung hình vảy 1 và mẫu bệnh phẩm ung thư biểu mô tế bào cổ tử cung hình vảy 2 là các mẫu bệnh phẩm thu từ các bệnh nhân khác nhau.

Các mẫu bệnh phẩm nêu trên phải trải qua quá trình giải trình tự bisulfite (BS) để phân tích gen đích (ZNF582) trong mỗi mẫu bệnh phẩm bị methyl hóa cao hay không. Kết quả được thể hiện trên Fig.1B chỉ ra rằng so với nhóm đối chứng, gen đích ZNF582 trong mẫu bệnh phẩm ung thư biểu mô tế bào cổ tử cung hình vảy 1 (SCC), mẫu bệnh phẩm ung thư biểu mô tế bào cổ tử cung hình vảy 2 (SCC) và dòng tế bào ung thư cổ tử cung HeLa (HeLa_0) đã được methyl hóa cao. Do đó, gen đích ZNF582 trong các mẫu bệnh phẩm ung thư cổ tử cung sẽ bị methyl hóa cao.

Gen đích: PDE8B

Các nhóm mẫu bệnh phẩm:

1. SiHa_0: dòng tế bào ung thư cổ tử cung SiHa;
2. SiHa_10D: dòng tế bào ung thư cổ tử cung SiHa (SiHa_0) được xử lý với chất ức chế ADN methyltransferaza 5'-aza-2'-deoxyxytidin 10 μ M (AZC);
3. SiHa_DT: dòng tế bào ung thư cổ tử cung SiHa (SiHa_0) được xử lý với cả chất ức chế ADN methyltransferaza 5'-aza-2'-deoxyxytidin 10 μ M (AZC) và TSA0,33 μ M;
4. SCC 1: mẫu bệnh phẩm ung thư biểu mô tế bào cổ tử cung hình vảy 1;
5. SCC 2: mẫu bệnh phẩm ung thư biểu mô tế bào cổ tử cung hình vảy 2;
6. Nhóm đối chứng (bình thường): ADN huyết ở cổ tử cung bình thường làm nhóm đối chứng chưa methyl hóa;

Trong đó các mẫu bệnh phẩm ung thư biểu mô tế bào cổ tử cung hình vảy 1 và 2 là các mẫu bệnh phẩm thu từ các bệnh nhân khác nhau.

Các mẫu bệnh phẩm nêu trên phải trải qua quá trình giải trình tự bisulfite (BS) để phân tích xem gen đích (PDE8B) trong mỗi mẫu bệnh phẩm bị methyl hóa cao hay không. Kết quả được thể hiện trên Fig.1C chỉ ra rằng so với nhóm đối chứng, gen đích PDE8B trong mẫu bệnh phẩm ung thư biểu mô tế bào cổ tử cung hình vảy (SCC 1), mẫu bệnh phẩm ung thư biểu mô tế bào cổ tử cung hình vảy 2 (SCC 2) và dòng tế bào ung thư cổ tử cung SiHa (SiHa_0) đã bị methyl hóa cao. Do đó, gen đích PDE8B trong các mẫu bệnh phẩm ung thư cổ

tử cung sẽ bị metyl hóa cao.

Gen đích: DBC1

Các nhóm mẫu bệnh phẩm:

1. SiHa_0: dòng tế bào ung thư cổ tử cung SiHa;
2. SiHa_10D: dòng tế bào ung thư cổ tử cung SiHa (SiHa_0) được xử lý với chất ức chế ADN methyltransferaza 5'-aza-2'-deoxyxytidin 10 μ M (AZC);
3. SiHa_DT: dòng tế bào ung thư cổ tử cung SiHa (SiHa_0) được xử lý với cả chất ức chế ADN methyltransferaza 5'-aza-2'-deoxyxytidin 10 μ M (AZC) và TSA 0,33 μ M;
4. SCC 1: mẫu bệnh phẩm ung thư biểu mô tế bào cổ tử cung hình vảy 1;
5. SCC 2: mẫu bệnh phẩm ung thư biểu mô tế bào cổ tử cung hình vảy 2;
6. Nhóm đối chứng (bình thường): ADN huyết ở cổ tử cung bình thường làm nhóm đối chứng chưa metyl hóa;

Trong đó mẫu bệnh phẩm ung thư biểu mô tế bào cổ tử cung hình vảy 1 và mẫu bệnh phẩm ung thư biểu mô tế bào cổ tử cung hình vảy 2 là các mẫu bệnh phẩm thu từ các bệnh nhân khác nhau.

Các mẫu bệnh phẩm nêu trên phải trải qua quá trình giải trình tự bisulfite (BS) để phân tích xem gen đích (DBC1) trong mỗi mẫu bệnh phẩm có được metyl hóa cao hay không. Kết quả được thể hiện trên Fig.1D chỉ ra rằng, so với nhóm đối chứng, gen đích DBC1 trong mẫu bệnh phẩm ung thư biểu mô tế bào cổ tử cung hình vảy 1 (SCC 1), mẫu bệnh phẩm ung thư biểu mô tế bào cổ tử cung hình vảy 2 (SCC 2) và dòng tế bào ung thư cổ tử cung SiHa (SiHa_0) bị metyl hóa cao. Do đó, gen đích DBC1 trong các mẫu bệnh phẩm ung thư cổ tử cung sẽ bị metyl hóa cao.

Ví dụ 4: Phân tích metyl hóa gen đích trong các mẫu bệnh phẩm ung thư cổ tử cung

Phản ứng chuỗi polymeaza metyl hóa đặc hiệu (MSP) được sử dụng để phân tích trạng thái metyl hóa của bốn gen đích trong các mẫu bệnh phẩm ung thư biểu mô tế bào cổ tử cung hình vảy (SCC). Kết quả phân tích trạng thái metyl hóa được thể hiện trong bảng 4 cho thấy, trong mẫu cổ tử cung thông thường, tỷ lệ metyl hóa của DBC1, PDE8B, PTPRR và ZNF582 lần lượt là 11%, 0%, 9% và 6%; trong mẫu bệnh phẩm ung thư biểu mô tế bào cổ tử cung hình vảy, tỷ lệ metyl hóa của DBC1, PDE8B, PTPRR và ZNF582 lần lượt là

100%, 47%, 100% và 97%. Cần được hiểu từ các kết quả này là, trong các mẫu bệnh phẩm ung thư biểu mô tế bào cổ tử cung hình vảy, bốn gen nói trên được methyl hóa cao. Do đó, tỷ lệ methyl hóa của DBC1, PDE8B, PTPRR và ZNF582 có thể được sử dụng thực ra là làm chỉ số sàng lọc trong sàng lọc ung thư cổ tử cung.

Bảng 4 Phân tích trạng thái methyl hóa của gen đích trong các mẫu bệnh phẩm ung thư biểu mô tế bào cổ tử cung hình vảy

	Tỷ lệ methyl hóa CGI của các gen đích (%)	
	Bình thường (n=54)	Mô SCC (n=30)
DBC1	11%	100%
PDE8B	0%	47%
PTPRR	9%	100%
ZNF582	6%	97%

Ví dụ 5: Phân tích sự methyl hóa gen đích trong mẫu bệnh phẩm khối u buồng trứng

Phản ứng chuỗi polymeaza methyl hóa đặc hiệu (MSP) được sử dụng để phân tích trạng thái methyl hóa của gen đích trong các mẫu bệnh phẩm khối u buồng trứng. Kết quả phân tích trạng thái methyl hóa được thể hiện trong bảng 5 thông báo các trạng thái methyl hóa của ba gen DBC1, PTPRR và ZNF582 trong mẫu bệnh phẩm khối u buồng trứng ác tính lần lượt là 50,3%, 50,0% và 56,3%, tỷ lệ methyl hóa của DBC1, PTPRR và ZNF582 trong mẫu bệnh phẩm khối u buồng trứng ban đầu lần lượt là 2,5%, 47%, 0,0% và 12,5%. Độ chênh lệch về mức độ methyl hóa lần lượt là 53,8%, 50,0% và 43,8%. Do đó, so với mẫu bệnh phẩm khối u buồng trứng ban đầu, các mức methyl hóa của ba gen này trong mẫu bệnh phẩm khối u buồng trứng ác tính có khả năng đánh dấu cao hơn. Do đó, tỷ lệ methyl hóa của DBC1, PTPRR và ZNF582 có thể được sử dụng thực ra là làm chỉ số sàng lọc trong sàng lọc ung thư buồng trứng.

Bảng 5 Phân tích trạng thái methyl hóa của gen đích trong các mẫu bệnh phẩm khối u buồng trứng

U buồng trứng ác tính	U buồng trứng ban đầu	Độ chênh lệch về
-----------------------	-----------------------	------------------

	(n=28)	đầu (n=28)	mức độ metyl hóa
DBC1	56,3%	2,5%	53,8%
PTPRR	50,0%	0,0%	50,0%
ZNF582	56,3%	12,5%	43,8%

Ví dụ 6: Phân tích sự metyl hóa của gen đích trong mẫu bệnh phẩm ung thư ruột kết

Phản ứng chuỗi polymeaza metyl hóa đặc hiệu (MSP) được sử dụng để phân tích trạng thái metyl hóa của gen đích trong các mẫu bệnh phẩm ung thư ruột kết. Kết quả phân tích trạng thái metyl hóa được thể hiện trong bảng 6 thông báo trạng thái metyl hóa của bốn gen, DBC1, PDE8B, PTPRR và ZNF582 trong mẫu bệnh phẩm ung thư ruột kết. Kết quả cho thấy rằng các tỷ lệ metyl hóa của DBC1, PDE8B, PTPRR và ZNF582 trong mẫu bệnh phẩm ung thư ruột kết lần lượt là 100,0%, 100,0%, 100,0% và 100,0%; trong khi tỷ lệ metyl hóa của DBC1, PDE8B, PTPRR và ZNF582 trong mẫu mô ruột kết bình thường lần lượt là 25,0%, 25,0%, 25,0% và 25,0%. Do đó, so với mẫu mô ruột kết bình thường, các mức metyl hóa của bốn gen này trong mẫu bệnh phẩm ung thư ruột kết có khả năng đánh dấu cao hơn. Do đó, tỷ lệ metyl hóa của DBC1, PTPRR và ZNF582 có thể được sử dụng thực ra là làm chỉ số sàng lọc trong sàng lọc ung thư ruột kết.

Bảng 6: Phân tích trạng thái metyl hóa của gen đích trong mẫu bệnh phẩm ung thư ruột kết

Tỷ lệ metyl hóa CGI các gen đích (%)		
	Mẫu mô ruột kết bình thường (n=24)	Mẫu bệnh phẩm ung thư ruột kết (n=20)
DBC1	25%	100%
PDE8B	25%	100%
PTPRR	25%	100%
ZNF582	25%	100%

Hiệu quả đạt được của sàng chế

Phương pháp sàng lọc ung thư theo sàng chế có các ưu điểm sau đây so với các giải

pháp thông thường đã được mô tả ở trên:

1. Phương pháp sàng lọc ung thư, chất đánh dấu sinh học và cách sử dụng chúng, theo sáng chế sử dụng tỷ lệ metyl hóa của một gen cụ thể trong mẫu bệnh phẩm phân lập được làm chất đánh dấu sinh học để xác định nguy cơ bị bệnh ung thư, và khi so sánh với các phương pháp thông thường chẳng hạn như nạo sinh thiết cổ tử cung và xét nghiệm HPV, cả độ nhạy và tính đặc hiệu của phương pháp sàng lọc ung thư theo sáng chế đều cao hơn so với hai phương pháp nêu trên.

2. Phương pháp sàng lọc ung thư, chất đánh dấu sinh học và cách sử dụng chúng, theo sáng chế có thể không chỉ được sử dụng để sàng lọc ung thư cổ tử cung dòng thứ nhất, mà còn được sử dụng kết hợp với hoặc hỗ trợ cho xét nghiệm HPV để sàng lọc ung thư cổ tử cung dòng thứ hai, để đạt được hiệu quả sàng lọc ung thư cổ tử cung chính xác hơn.

3. Phương pháp sàng lọc ung thư, chất đánh dấu sinh học và cách sử dụng chúng, theo sáng chế có thể được áp dụng không chỉ cho sàng lọc ung thư cổ tử cung mà còn cho sàng lọc các loại ung thư khác (ví dụ: ung thư buồng trứng, ung thư ruột kết và tương tự) để hỗ trợ chẩn đoán ở mẫu bệnh phẩm bất thường.

Tất nhiên, có thể thực hiện nhiều thay đổi và sửa đổi trong phương án được mô tả ở trên theo sáng chế mà không tách khỏi phạm vi của sáng chế. Theo đó, để đẩy mạnh phát triển khoa học và các lĩnh vực sử dụng, sáng chế được bộc lộ và được dự định chỉ được giới hạn ở phạm vi của các điểm yêu cầu bảo hộ kèm theo.

YÊU CẦU BẢO HỘ

1. Phương pháp sàng lọc ung thư, dựa vào việc phát hiện tỷ lệ metyl hóa của gen đích trong mẫu bệnh phẩm, và sử dụng tỷ lệ metyl hóa làm chỉ số sàng lọc để xác định nguy cơ mắc bệnh ung thư, phương pháp bao gồm các bước sau:

bước 1: chuẩn bị mẫu bệnh phẩm;

bước 2: phát hiện tỷ lệ metyl hóa của trình tự CpG trong gen đích là ZNF582; và

bước 3: xác định liệu có ung thư hoặc thay đổi bệnh lý tiền ung thư hoặc không trong mẫu bệnh phẩm dựa vào tỷ lệ metyl hóa trong gen đích, hoặc sử dụng tỷ lệ metyl hóa của gen đích làm dấu ấn chẩn đoán.

2. Phương pháp sàng lọc ung thư theo điểm 1, trong đó mẫu bệnh phẩm là mẫu bệnh phẩm được phân lập chọn từ nhóm gồm có mẫu sinh thiết cổ tử cung, dịch cổ tử cung, máu, nước tiểu, phân, nước bọt, tế bào màng nhầy ở miệng, dịch dạ dày, dịch mật, tế bào biểu mô cổ tử cung, và mô ung thư hậu phẫu.

3. Phương pháp sàng lọc ung thư theo điểm 1, trong đó phương pháp xét nghiệm tỷ lệ metyl hóa của trình tự CpG của gen đích bao gồm phản ứng chuỗi polymeaza đặc hiệu metyl hóa (methylation-specific polymerase chain reaction-MSP), phản ứng chuỗi polymeaza đặc hiệu metyl hóa định lượng (quantitative methylation-specific polymerase chain reaction-QMSP), giải trình tự bisunfit (bisulfite sequencing-BS), vi dây, khối phổ, sắc ký lỏng hiệu năng cao biến tính (denaturing high-performance liquid chromatography-DHPLC); và giải trình tự ADN dựa trên sự giải phóng của pyrophosphat và kết hợp bất kỳ các phương pháp trên.

4. Phương pháp sàng lọc ung thư theo điểm 1, trong đó gen đích ZNF582 có trình tự nucleotit được nêu trong SEQ ID No. 4.

5. Phương pháp sàng lọc ung thư theo điểm 4, trong đó giải trình tự bisulfit metyl hóa của gen đích ZNF582 được thực hiện bằng cách sử dụng cặp mồi có trình tự nucleotit được nêu trong SEQ ID No. 19 và 20.

6. Phương pháp sàng lọc ung thư theo điểm 1, trong đó gen đích ZNF582 được nhận biết bằng cặp mồi có trình tự nucleotit được nêu trong SEQ ID No. 11 và 12.

7. Phương pháp sàng lọc ung thư theo điểm bất kỳ trong số các điểm từ 1 đến 6, trong

đó bệnh ung thư là bệnh ung thư cổ tử cung, bệnh ung thư ruột kết hoặc bệnh ung thư buồng trứng.

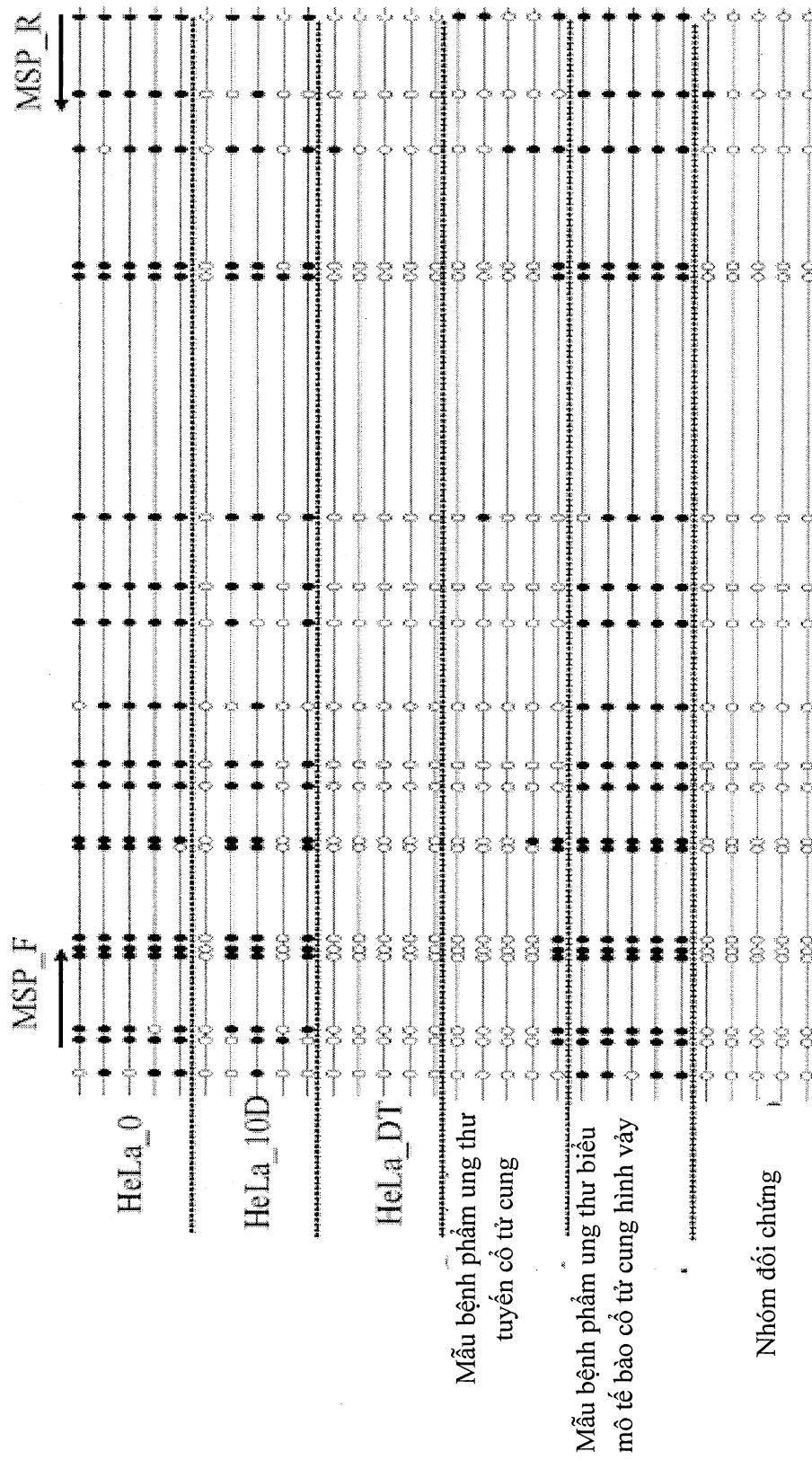


FIG. 1A

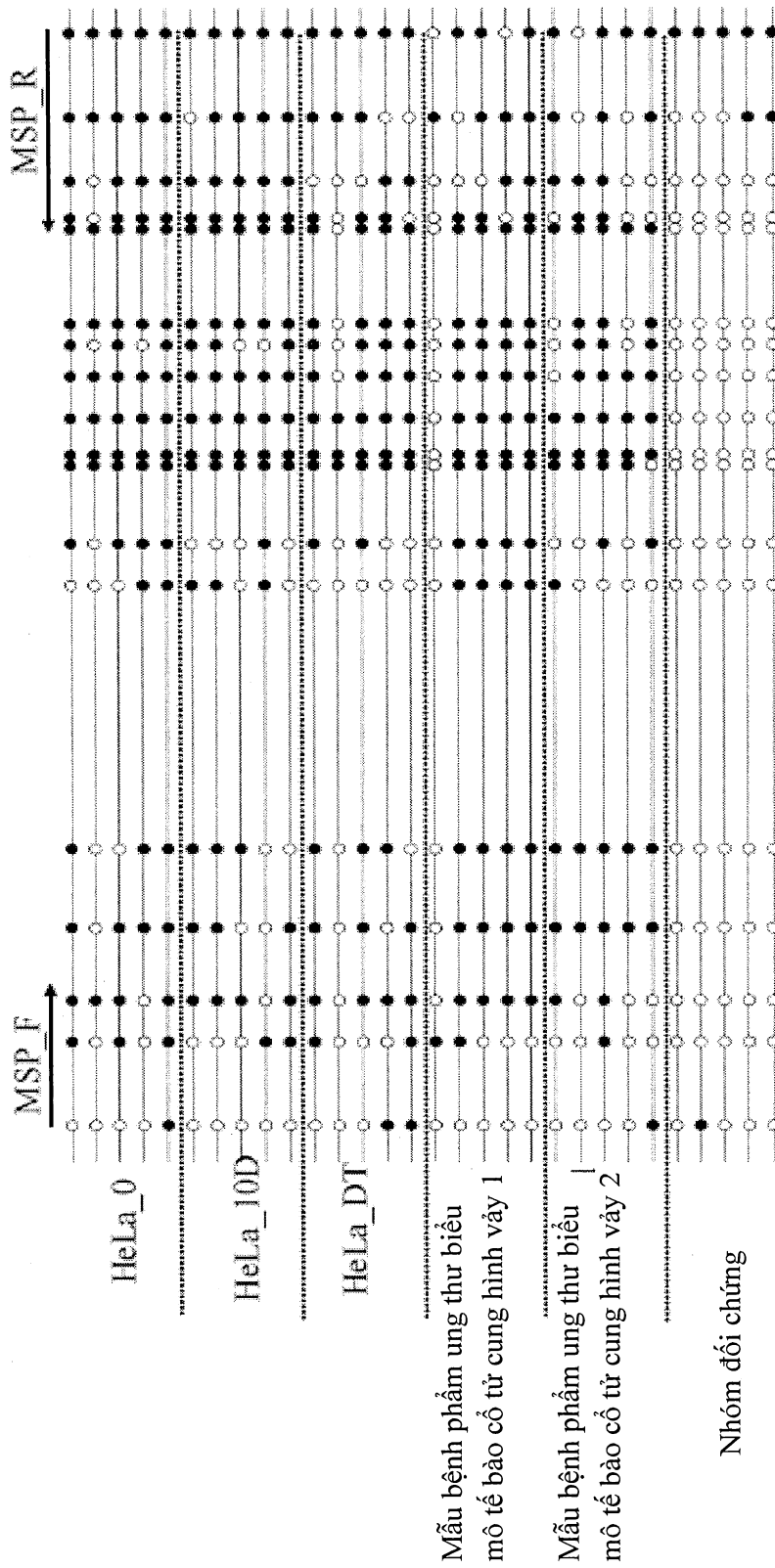


FIG. 1B

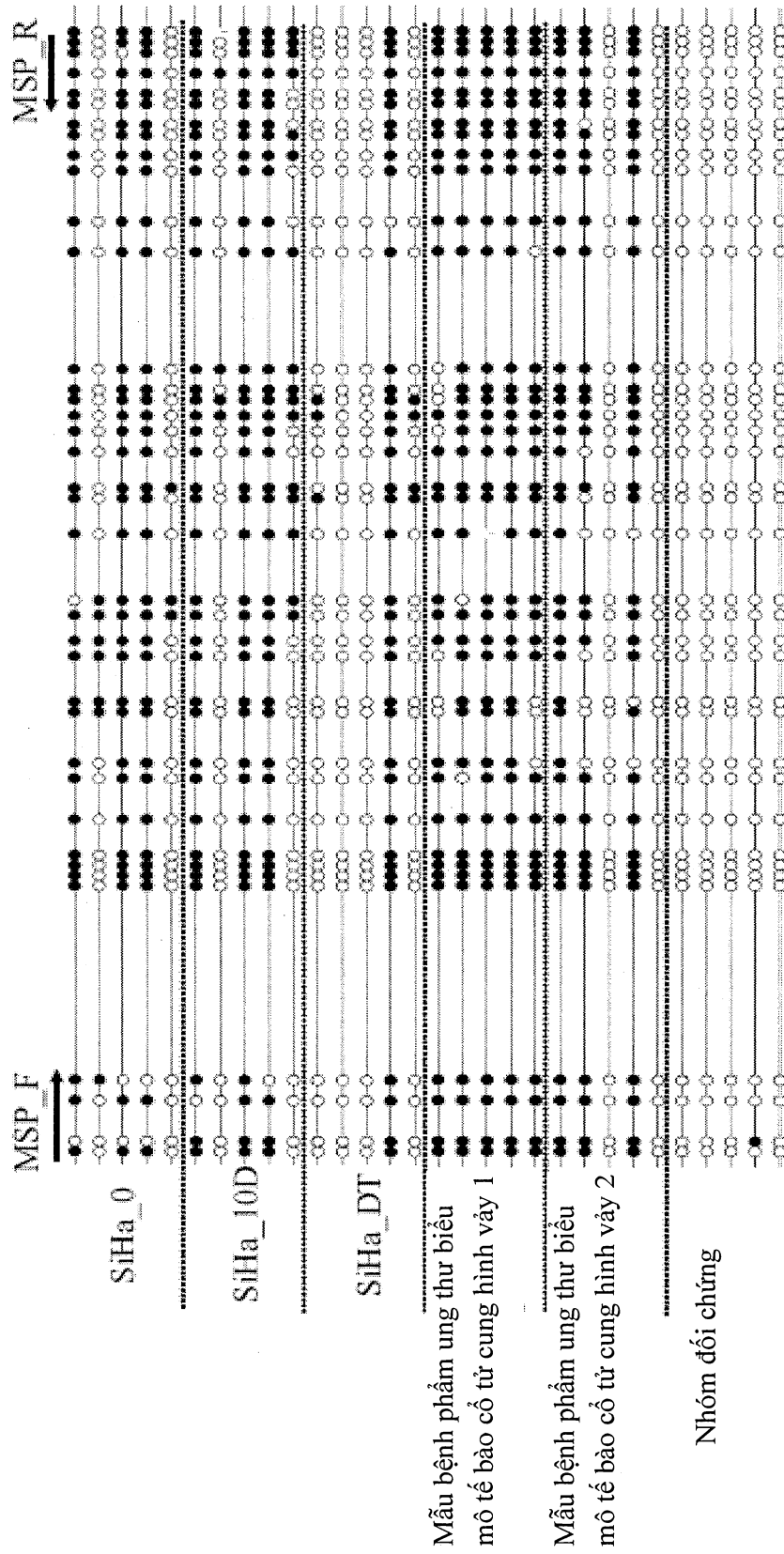


FIG. 1C

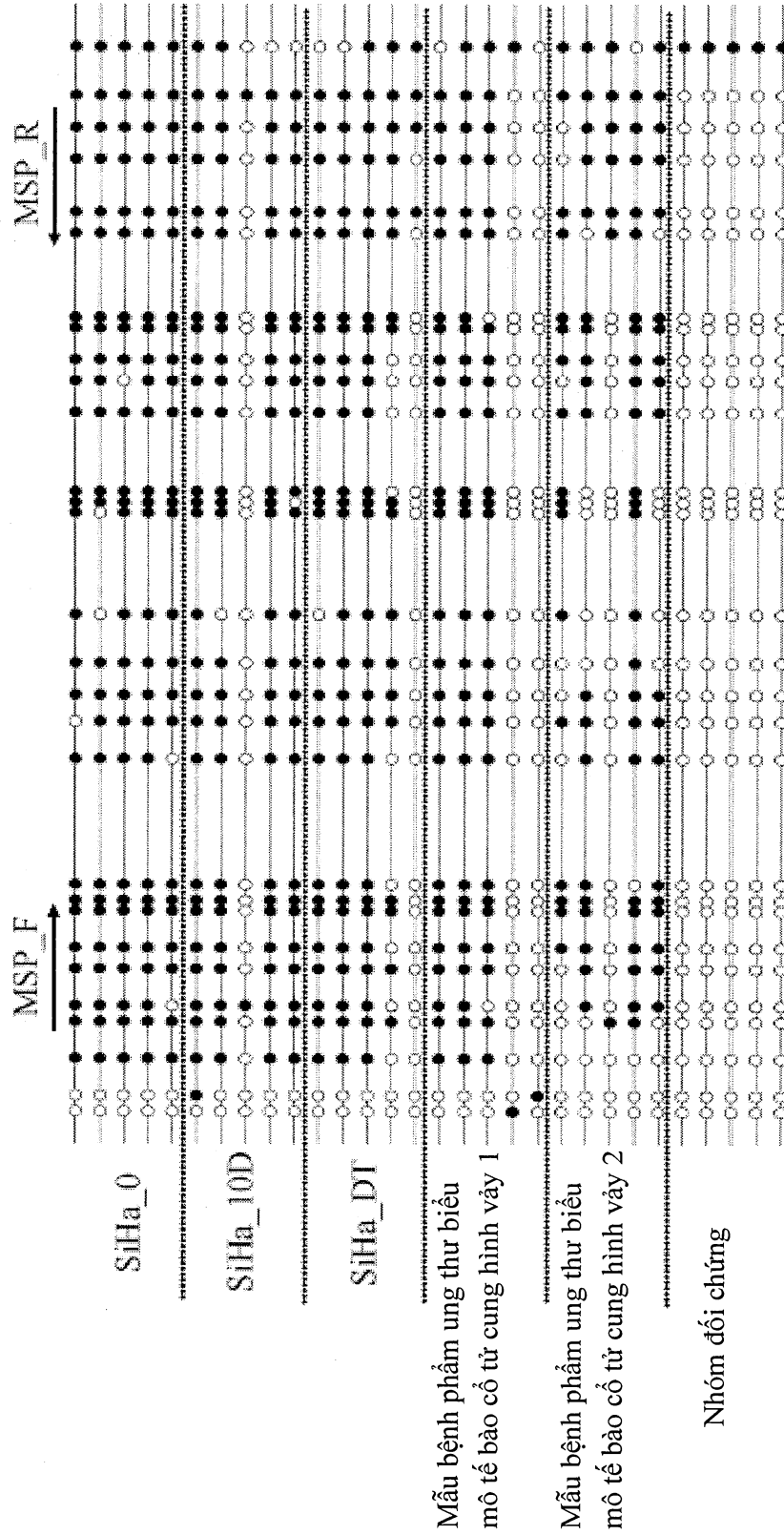


FIG. 1D

DANH MỤC TRÌNH TỰ

<110> Hung-Cheng Lai

<120> Phương pháp sàng lọc ung thư

<130>

<140>

<141>

<160> 24

<210> 1

<211> 5000

<212> ADN

<213> Homo sapiens

<400> 1

ggtcttcacc tgaagttgcc tgactacgtt gtctcagatt gatttcagaa ctgcattatc	60
agcagggcag ctgttgcceta ctaacagtca tggaagtaga tctcagctgg gggatgatgg	120
ggaaacaaag agaaagatag ggacgaataa aggaaggaat ggggcacaga cagaaagaga	180
gagagagaaa agaaagaaag gaggagaggc aagagatagt gacctggggg ttcataccgc	240
ggtgggtggg agaaggagtg tgtgtgggtg gcgtatttct tcgtggtcag aatcaaagtc	300
agaactgacc tggggaacct gtgttcaaga cctgaacctg gaccagaagc caagggaaag	360
agaaagttgc ccaagaaaaa agggagaaat ttccaccagg cgaaaataga aatcgctgac	420
tgggtttgtg gctggagagc tgtccgtggt gctgactcct cctcattggg attccagtga	480
aggctgagaa gctctgcact ggctgctcct cctctccctt tcctctcccc tcagtcctga	540
ccttctgaga cccaggatca ttctggctag ctcgatgtct tctcctctct ttctcccttt	600
ctctttgtcg attaaataat ttttcccgag ttcccagttc atcactcaaa catttctctt	660
ccattttgtt tcaagactcc tcttctcca attagttcac tggctcctag attgttctag	720
gcctccctgg tggtcacatc tctctcttat cctcctttga attaaaaaac aaacaaaaaa	780
catacaatgg caatcataaa agcagacaac atgcactgag cactcactag gtgtcaggga	840
gcagtetaag tgcctttgca tacattaact catgagcagt ctaagtgcct ttgcatatat	900
taactcatta caattccaca tcaaccctat tagatgggta gatataattat tattctatta	960
aacagacaag gaattgagcg ttaaggggta agtgatttca ccaagaccag acagctacag	1020
gactccaacc tctctggagc ctgcacttct actgaatctg gctcttttac atattccaac	1080
accttttct atgacttctt gccctcccat ctggtagcct taatgctttc tattcttgcc	1140
taatcttcag gggactaggg atctgcattg ctgatttggt ttaaatcaga agggaagcaa	1200
agtaaacaaa cataccaaat gtgcgctgtg tgggaattat catacagagg ctttattttc	1260
tgcttcagga agaggcccta tgtagcagc cccagcctgc attcaggctg attgcagagt	1320
atcttgcttt ttattttcat gtcttagtcc ctgtaccctc gcccttccc cgctcttggt	1380
ggtctccaga gaacttcgtg tcccctcagc ttctccctcc tacatcctgc ctacgtagag	1440
aagctcttgc ttcattctgg gaggttacgt gggtctctgc ctacacaccg agagaaacaa	1500
acagtgtcaa aactcacag agagacgcgc agacacaaac ggaccacac gggcaactcc	1560

cgagacaaaa	cccacactcg	atggatccac	gcggccgtgg	aaacacctgc	cgccccagaa	1620
acactcaggt	actcgcgaca	cacacagtac	agtcacgctt	aagggcacca	ggattccggg	1680
tttgcgcgta	tgcgcgggtcc	ctttggatgc	tcgtgcgcat	agacacaaca	ccctacacgc	1740
cccagaccca	cgaaactccc	tacggctcag	ccccagccca	cccgggcccgc	ccttccctcg	1800
aggcggcctc	ccgtctctcc	tcctctcgct	tctcctcctc	ctccgcctaa	agatgtacaa	1860
aacactcctc	ggaagcaacc	ccggcggttc	gtcctcctct	ccccgcccc	cgcccgccgc	1920
tccccattc	atcttcggcc	gtcgcgggct	aagtccctcc	cccggcgtag	cccggcctcc	1980
gccgtcccc	gcccgagag	cgcggcgcac	ttggacttcc	ctctccattc	gccagccgcc	2040
tcgtccccgg	acccacggc	tgcaaaactga	tctggcgcgc	ggggaggagg	agagcgcagg	2100
cgagcgaacc	cgcgagagag	ggagagagcg	agcagacaa	agcagagagc	agagcagag	2160
agccgggagg	cagagggagt	agtgaccgcc	ttccggagcc	gggattcatg	cctgtcctcg	2220
ggaccagcga	aggggacttt	acggctgagt	atgagccagg	ctgctaggag	ccaggtaccc	2280
ccacgcctgc	agtcctccgc	ccgtgcccg	aatgcgagct	gcacgcaggg	ctctcccaag	2340
ttcccaccga	gccgaataaa	aagcgtcctc	ccgcagctct	ccgccaaaga	cggacattga	2400
ctccaggtaa	ggcggcgccg	ggtgcagcgc	cccgcagccc	cgtgcctt	ggaccgcgcc	2460
ccgggcccga	ttcggggcgt	ccccgcgctc	ctctgcccct	ccccctaccg	gcacccttgt	2520
ccgtctttca	cctggccgcc	ccgccgcctc	caagtcttct	ccagttctag	ggaggggggt	2580
cctgtgcctg	gggctcaaag	ggctaattgc	gggtttgagt	gagtggcgtg	tgtgtccccg	2640
cgcgctcccg	acgtgtgcac	catggtggga	acttgatgtg	gtgctagtgt	gtttgcgtgt	2700
gcggcgctgt	ttgcgtttga	tgcgtgtgtt	cgtggtgtgt	gtgtgtgtcc	gtgtgtgtaa	2760
gggaggggtg	aagagagaga	ggtcctataa	cctacttacg	gcgcgatgtg	tgtgtgcatg	2820
tttgtacgta	tgtgtttgta	tgtgtgccct	cgtgtgtctt	tttaattagg	tctctccagc	2880
ttacacggaa	tgggaccctt	actataggat	cacgtagtca	ccgggaaacc	cgtgtggac	2940
ttcctcttgg	ggctctgggc	ttggggtttg	gggaggatta	tggggctgta	gatggcacct	3000
tatttagccc	aatgtttgta	cgtttgaagg	aaaattcctc	caaacgggtg	aatcctgcta	3060
cactgggacc	cacagcttaa	tatgaaagag	acatggccaa	ccccgaggc	aaatgagacg	3120
ctgtcacttt	aaattctaca	ctgggcagac	tccaagattc	tgatgggaat	ttggagacac	3180
tggatagtgg	gtgacagaga	aagggggaag	tcagcgggtg	gtccttatac	tggggccttg	3240
aaagtctggg	atagggattt	accctgcac	ccggtttgca	atcaacagag	cccctccggc	3300
ttctgtttgt	tttggggagg	atctggcaag	tttgcattga	tccccctcag	gggaaaggag	3360
aaagcgtcct	cggggacttg	ctcatccatc	acagttgcaa	agggtctcag	aggaaatttc	3420
atctggggcg	gctgtggatg	atatggaagg	agatggatgg	ggcactttcc	taagacagat	3480
tcgtctttct	ttccccattt	caggcgggga	agccccaga	gagtctctcc	ctaaatatgc	3540
ctctccatgg	ctccccctga	gttaggggga	tattgagaga	aggcagagag	gtggagaggg	3600
agagagagag	agcaagagcg	agagagaggg	agacagagag	agagagtgtt	tcagtactga	3660
gggagatcta	caatttgaaa	aggggctgtg	agtgtggaac	ccatgacaga	atgtggcagt	3720
aattgactta	aactgctgtc	ggtttgcacc	tcgcgtctct	cacttggctc	caaaatactg	3780
ccaagggcag	gggggcgggtg	gaggaatctc	agccaggaag	gtagtttggg	tcaaggccct	3840
ctttcttctt	tcctccatct	cctgggaggc	tcttttaggt	ggtccttgg	ccaccctggc	3900
ttggtactgt	tgggcttggg	ctctggggcc	agccattgat	ttgattccca	gctgcctctt	3960
aaaggcttgc	cctactcagc	aaaaatgctg	agttttactg	ctgttagcat	ggcttccaga	4020
ctcggcagct	gttactttct	caaggtaagc	ggcagccgtt	ctgctctcag	gcagggaggc	4080
taggagaaga	caagggtgtg	gacgctggga	cacagcctct	tgtctagctc	agcacggagg	4140
ggccgctgaa	aagccctctt	aggggaaaaa	agtagaaata	tatttggcct	aaaaaatgta	4200

cacatatatt	ctaggaagat	atatatatat	atatatatat	caatcacaca	cacacacaca	4260
cacacacatg	tgttcatatt	ctggggagag	ataaaagcaa	acatgtattt	gtaggaacct	4320
gccctagtgg	gtgggttcta	cctgcaggca	cttgcagaca	taccattctc	tgtccagagg	4380
ctgatacctc	aggctgaagc	cttcacatag	ccacagagaa	tctccttcag	gaaactcatg	4440
tcagggcaga	gctccctgga	ttatcttggg	taggggttac	tagggccaag	gaggagaccc	4500
atTTTTGGGA	actgcctcat	tttctctctt	tcctaattca	cagacgatcg	aaactggaag	4560
gaatattaaa	gagcagaaag	gcaggactct	gtggcctaaa	gagggaatga	gaattcccca	4620
acattcagta	gtgggttagt	ggcatgccca	ggactagaat	ctaggcttat	caatcagcta	4680
gatcaatacc	ctttccactc	ctcccagcaa	tatacactct	aggactgtat	attgctccca	4740
cccctatctt	gccaacctgc	caccatgacc	cctaccaaga	caatcgcttg	tcttccctca	4800
tcaaacttca	cattttcaga	gatgcatttc	agggtttttc	tatctggaaa	catggccctg	4860
aagagaaaag	atccaatgcc	tcaatcacta	ttcccgaag	acacacacac	acacacacac	4920
acacacacac	acacacacac	acacacacac	acacacacgg	tggtgggggg	aggactgaga	4980
gagagacaga	gagagaacat					5000

<210> 2

<211> 5000

<212> ADN

<213> Homo sapiens

<400> 2

agaacaactc	tgaagcatca	tcccaccttc	agaatcctcc	ctgcagggat	gcctgagget	60
tggtcaattt	ctccctctgc	tcaagcttgc	ctccctccat	tccctccctt	tatttcctcc	120
cacaaatgtt	gactacaaga	gcacttccta	aaaagcctcc	ctctgcctgc	taaacacccat	180
ctcagaatct	gcttctagga	ggacceaac	tgtgcagaaa	ttttatgagc	aatttgccctg	240
tccattaaaa	aatacagcat	agactttttc	caggtttagct	ggagatggag	tttttttgta	300
tcaaaaaaag	agagctaaac	tgagaagaga	agcaaaggta	gaagaataag	gttagactct	360
ccctgtataa	ctgaatttgg	gattgggaat	gatatcttgc	ttctctatga	aatttagaaa	420
ctgtcttttg	aagtcttcat	ccagtcacat	ccaagagtgg	tccaggaaga	aaaaggaaaa	480
taagaatgaa	aaaaaaacaa	aaaacaaacc	ttaaaattgg	atacttcttt	gtgccttgca	540
ttaagcctct	atgttactgc	ccttgctcgt	ataagcaact	ctagataaaa	tcctttctta	600
gaccatctaa	tttatctccc	cgtgccattg	tgaaaatcgg	gcaggaatta	ctgcgggaac	660
cagtactga	atgggcagtg	gtgtctcgca	gctgtaactc	agagtttaca	agggaaggat	720
gaacagcagc	catccttcta	aaggcattct	tcctctcctc	cagcacagct	cagtctgtgc	780
atgcctcacc	aaggagaatg	agggaaggcc	ttgttgactg	cccagcccag	agtgtgagga	840
acagcagttg	acatgggcgg	ctggacggga	tagaaaccag	catgaacatt	tagtgacacc	900
taaggaagag	gtgaaaatgg	gggagaaagc	caagaacat	cagacagctt	ctttctgttc	960
ttgccttgga	tcttagggga	acgtatgtgt	tcaagggtgc	cttctccgc	agggcctccc	1020
atcccttggc	acaaacttcc	aaacgatgag	tagacatgga	atgggcttct	tctgaagact	1080
gggtcaccat	ccaactcagc	cttcggagta	ggcgggtggg	ctgggggaag	aggacggaac	1140
catgaggtct	ctccatcact	tcccgggtgt	gttgagaggc	aagcaaatgc	ggtgggtggg	1200
cctgggctgt	agggccctg	ccatttcggg	aggggcctcc	tggtgtttag	ctctgggatg	1260
tgtgaaaatg	tgttggtgaa	aagcgagagg	cttacacagc	ccttcagggg	aagaggggct	1320

ggggcgctgg	gggcggcgctc	gggatgagtg	cagaagagac	gaggcctctg	gacagcggag	1380
gaggagggga	gggcgcccgag	gcgcggtgcc	agctgcccg	cacaggggcc	ccgcggcgga	1440
gccgagccgc	gggcacgctc	tgcctgtcg	ggagagctcc	gggagcggcg	ggagggcg	1500
aggggcgag	tggggcccg	gcggctgcg	ccgcggagcc	ggggcacctg	aggaggaagg	1560
agggtgggag	cgagggaggg	aggggacggg	cgagaccga	aagtggggaa	agaaggtgca	1620
ggcaggcggg	caggcggcg	ggcgccctgg	cccaggccg	cgggtgcggg	agcccggcga	1680
ggtcgagctg	ggcggcgcg	ggggccgcgc	cgagggagga	ggggaaggcg	gaggcgcggg	1740
gagcgtgttt	ggggcgccgc	ggcggggagg	gtggcgccg	ctggtgcgcg	cggggcgctg	1800
tgtatgcgcg	ctccccgct	cggggaggaa	gatggcccaa	aagggaagt	tggggtgacg	1860
cgcgcggtcc	ccggaggctc	ggcggggggc	accgcggcca	gcccagcgga	gcggcgga	1920
cacaggccgg	ggggcgcgca	gtccgggcgc	cgcccgggc	gccccctcac	tgcaggtggc	1980
agcgggtgcg	ctgggtccc	gcggccgcg	gcgcggcg	gcgcgcggg	gagcccggcc	2040
gagggatggg	ctgcgcccc	agcatccatg	tctcgagag	cggcgtgatc	tactgccggg	2100
actcgagca	gtccagctcg	ccccgccaga	ccaccagcgt	gtcgagggc	ccggcgga	2160
ccctgcccgg	cctcttcgtc	cagaccgacg	ccgccgacgc	catcccccg	agccgcgcgt	2220
cgggaccccc	cagcgtagcc	cgcgtccgca	gggcccgcac	cgagctgggc	agcggtagca	2280
gcgcgggttc	cgcagcccc	gccgcgacca	ccagcagggg	ccggaggcgc	cactgctgca	2340
gcagcgccga	ggccgagact	cagacctgct	acaccagcgt	gaaggtaa	gccccgcgct	2400
ggcacacgcc	gtggggggcg	tccgcccgt	cggcggggct	cgcacgggta	gggggctccg	2460
gcggagttag	gtgaccgtga	ggcggttggt	ttggagaggt	tgtcactaag	gaggagttaa	2520
cttttcattt	gtggagatga	tgggagccca	ggaaatgtgg	tcagaaaaag	gccccgtgag	2580
gggtcctgga	agcgtcctta	gctggtcctg	ggggactggg	cggggaagg	agcgcagaag	2640
gaagcaggtg	ggctggcctg	ttcctccttg	agggcaggaa	ggctgtggct	tggtttatgc	2700
aggaagaggg	gtggggacca	ttgagagcat	tccgtggcca	gtcctgttga	atgaaatctg	2760
agcactgagc	tggatttgcg	tgcctttag	gtactggtg	cagttgcagc	accaggatag	2820
atagtcccc	atattccgat	ttttacctgg	gattaccagc	caggctggag	tctcagcaca	2880
ggaaccgagc	gtagggattt	gtaatgaat	gagtgttcgt	gttttaagag	atgtgggaac	2940
ggagcagagt	ggaacctgtt	gtttgtcact	gtaacgtttc	tctgggttgg	ctgcatccta	3000
gacagaattg	agagaaccgg	gccatgagtt	cggagtgtca	gcagagccac	cgtgagggga	3060
cgtggtttcc	agtgcagtac	agctgtctga	ggatgattct	gcacatacaa	ctgatcttcc	3120
cagagagtgg	gattttgagg	aagtggaaaa	aattgttttag	atggttaaag	cagtcgctaa	3180
atatttattt	cttaaagaca	gaagaaaaat	aattatttaa	atagtgtcct	cccgtatgtt	3240
ctcaaagtac	cgtaaatacc	aagaggcttt	tcattgtgta	aatctgggca	ctgggtcttt	3300
tttcctttca	gcaaacaaga	taacaatggc	atatcctatt	gtacagagag	aaaaaaatac	3360
tctaattgtc	agatagaatc	acagccttta	cctgggtcaat	aggttacca	gaaccattcc	3420
atggattatt	tgtcaaagac	acatttgtat	ttttaatagt	taaaaacttg	gtcttcattg	3480
gatgaaggca	gcccacagtg	gaggagttag	gagaaggata	acatgttttg	gcttcaatct	3540
ggaagtccaa	aagtttttta	atgaacctat	ttttttttta	acagcatctg	attgtttaac	3600
atgaagcttg	actgatgttt	gcttgtgggt	gcagtgtttc	tctggcagta	atcataatca	3660
ccattatagt	aaaaacatcg	tttattgagc	acctactatg	tgccaggagt	aagctttctc	3720
tcctaactat	tgaatgttaa	cagtgtgatc	ttagcatagc	atgtctgaag	actagagtct	3780
aggatttatg	acagtacaaa	cctaccagtt	gcccattcgt	tacagaggca	tgtgctgaac	3840
gctatacttt	gcttttgc	agctctctct	ctctccccct	ccctgttcac	acaaagttag	3900
ctttaggagc	aattttcagg	gatgaaaatg	ttaaatttgg	tgaataatatt	aagttgggat	3960

attatcagta	aagaccatcc	ttttgcctga	cggttgaaga	cctgaagcag	tagttcaaaa	4020
ttgtgacagc	tcatgaactt	tgtgatactt	tttttggttg	ctccagagaa	attgtgaatt	4080
tttgttttct	agaatatcta	taatctcagt	ttagtacttg	gcaactctag	ctactccatg	4140
tttgttcagt	agaagaaaga	atactgaatg	aaccattctt	gtacattttt	aatacttccc	4200
tcaagaggct	gcttagagac	caaatgaagt	aaaaactgaa	gggacaatga	gccccaatga	4260
ttagcagctt	gtataaactt	gtattaatct	cattgctact	tggcttaagt	attaaaaggc	4320
ttgacctggg	taggagttta	actcattttg	ctttatcaag	aagattccca	aataagagct	4380
taattcactc	taaccatta	ctgggaatgt	tggaaatgaa	acccttggtt	cattgactta	4440
acactacttt	gcattgtgtg	gatctcagaa	agtatatatt	tgaagtgaaa	tttgagacca	4500
gagttattta	tgcaaatctt	acagaattat	tatcacagaa	ctgccctaag	atattttaata	4560
gtatcctttt	aggcctgtga	ctcgggtgtg	tattgtcttt	tctcaatcgt	aaagttcaca	4620
tgttctgtgc	ccttttggtg	ctgagataat	tttataaaaa	ctttacattc	atttttttgt	4680
agactaagca	ttcttttttt	gccgtaagag	tttgagatag	ttgctaggca	cgtgggtggg	4740
ccccaccct	ttccaaagca	cacaggtgga	aaaataaaat	ctttaccatt	tggctcctgtt	4800
tacagtgggtg	gaaggttgta	atagtttcca	ccccctcccc	ctggccattt	tacatcaatt	4860
atgaagtgtg	tcctcattat	ctacttggcc	tctggtattc	cccagagtgc	aaaaatgtgt	4920
gcttgatctc	tctctctctc	ttttctagca	ataattcatc	tcattaagct	tttagaatga	4980
tgaggcatta	tgatgtcaaa					5000

<210> 3

<211> 5000

<212> ADN

<213> Homo sapiens

<400> 3

aagaaaaata	gacctattgt	aagtgggata	caattatttt	tccctaacct	ccagccctgc	60
agcattttcta	gtaacagatg	ttagatgaga	agagttattc	ttctaagctt	tacagaatgc	120
taaataaatt	atctcaaagg	ttgactcagt	aagaatttca	aagagttaat	tgtcagtatt	180
gaactttgtc	aactatcatt	ttcattttatt	catttttgag	tatctatttt	gcatgaggca	240
ttgctctagg	gagttgggac	atatcaataa	agtcatacaa	gttcttcaac	attggagctc	300
aaattctgtt	ttgcgggggg	gagtagggaa	atagatacta	aacaatagat	tttataagta	360
attattttgg	tttgttagaa	agtgatacgg	gaaaagcaaa	ctggaggaca	gtgaatcagg	420
aataggagaa	gaggttgcag	cttagagtgg	tttgggatga	aatgatagtg	ccacagaaat	480
ggagtgtgcc	acacgcatgg	gggagtaggg	ggtggagcac	gttcttagag	aggggctaga	540
catggagata	tgtctctatt	ttcagcagag	aaggcctgtc	aagatattaa	aagtacatct	600
cattgtcccc	ttctcccact	tgttggagtg	cttcttgccc	ttgtctctca	gagctgaaga	660
tgacataaat	cataaaagaa	ataaatggta	ccgcctgggc	tctctgatct	tataaggtcc	720
cacctaattt	aagctctggt	ggatagccta	tttcccatta	gttgtctcaa	ttcacagagg	780
atggagaaaag	aaagcactga	atgagatgtg	cagctcaact	caatgggcac	ttgttagccc	840
caattatttg	gctaggacaa	ctttattcaa	gttaatggaa	tacaaaactc	agcagttctt	900
atggagactt	caatgtgctg	tagaaciaaac	cacagataaa	tgaactgtct	gaaagaataa	960
tatgaggcac	ggctaaacaa	gtgttgtgga	agagaaagat	cactgtaatc	tagagttagt	1020
tagagttagg	gctttctgga	aaagtgagac	caacattaca	ccttcgaaaa	ctagtgtggt	1080

ggtggtgcaa	gtggaggaaa	atgggaaaga	ggcaacagt	tccataaaga	cacagaggca	1140
gaaagggcag	aggccctggg	atagttaaaa	gagaagtcta	gctgaggggac	ttcttgactt	1200
gactagcagc	aattagccat	gactaataag	gctttccaca	ctccaaagac	ttagatggag	1260
ggataaaaaa	ccatctaata	gcagactgtg	gtagcctccc	tagagacaca	gagctggggc	1320
ggatgagtcc	aggcactgac	gtgatccatt	atctttcacc	ttaaagagta	aaagggaaac	1380
taaagttaat	tacctccacg	aaacaaaaag	gtgccttctt	gtgcttcaat	tacatggata	1440
tattctacta	gtctaaaagt	atcttctcac	ttctttctgt	cactgtgagg	acttgagtca	1500
gaagaaagtt	taaatacagt	cattgagctg	gaaagagtgg	aaagagaagc	aaagaggggg	1560
aagctgtagg	aaggacgaag	tcacccccaa	gatacatggg	tactgcttac	accaagcaag	1620
ctgccttggg	aacgcttccc	ccgagcagcc	agaatgctca	gcagtgggaag	acacctctat	1680
tcctgtaggc	gagtcctggg	aagctgggtca	atctgcaaat	gccaattccc	agcagtgagc	1740
tcggtcacg	tgtaaataca	gatttgggga	aagagtaggg	tgggtggcat	ggttgacaat	1800
gtcatcagct	ccctcctctg	actcctgtgg	tcgtgcccc	atctactctc	actcagctac	1860
accccacctt	cggatttgtg	atggacgctg	ggtccttagt	aaccacagca	agtgtctccc	1920
ccgcacttcc	cccttcccca	ccccacccc	cacccccaac	caccaccca	gcgatggagc	1980
ctactctgct	ccaagccgcc	gctaagaccc	ggagaagcgg	aatttcaact	tgaaattccc	2040
ttgcctcgtg	agggccggcg	ctgggcatgc	tcagtagccg	cggcgctgct	gctgggctgc	2100
tgggctggcg	cggagtcac	cctgccgtct	ccgccttggc	ttctgggcgt	ccagaaggcc	2160
aggcatttgc	cgcctctgag	cgttctgttt	ccccttacc	gcaacctcct	actgctcttc	2220
ctctctccct	ctcttaggga	ggttgaagct	ggtgctgggt	tctgtcggcg	ccacagactg	2280
actgctctgc	aaaccccagc	cgaggacctg	aatcccggag	actagaagac	ccttggcggt	2340
ggctctttct	aatagcactt	tacctgaagt	ggggctcgtg	tggagtttct	cctccacctc	2400
tcaatgcaaa	cactatgcgg	agagcagctt	gcttccctgc	gctgtgcctg	ctccttaatc	2460
ttcacgctgc	aggtaagggg	ttgccaggct	tagagccgga	gctgtgcatg	agatgggaaa	2520
ctgcacatgc	ttaaggactc	taggaaaagc	ttgccttgcg	aagaaagcct	ttctaaaaag	2580
gtaaaacagg	acagtacctg	acagggaagg	ggtgagggag	tcgtgcacct	gttggaatac	2640
tgaggccgaa	aacttaacct	aaaattagct	cttgattttt	ctttacttta	tacgaaaata	2700
gtgaacattt	tcaatatgaa	tacaggtttg	ctccccctt	cgcctcccc	ccgtctagat	2760
tccttgcctt	tgtttctagg	ctatgcactt	aactgtcaat	ttttcaggag	agggataaga	2820
catcctgcta	gatgtaacct	tttctactgc	agcggctact	acattcataa	ggtctcttgg	2880
tctagcgagc	gctcatagga	aggcattggc	tgtaacctga	tggaccacat	ctccgcccac	2940
aagatcgaaa	tacatggttc	acattagtg	tgctgcacca	ggctcctttg	gccccgtctc	3000
ttgtcactcg	aattttctaa	gccaataaga	ggaagaaaag	gtttcaaata	gaatcctctt	3060
gtctctttca	gcaccgtggc	tagcgtccgc	tggcatatta	aaaaaaaaaa	aaaaaaaaaa	3120
aaaaaaaaaa	acaattacag	ttccaaagct	gacaacctg	ggttaggcac	tgccctcg	3180
acagatttga	taaggtgtaa	ccaagagtaa	cagtgcagct	gctcccaaat	ataaatagat	3240
ttgaggcacc	agaatcaata	tcaagtctcg	tgaggaagca	aaccttggga	agtacctgag	3300
agatattata	tgtggtgtcc	ctgctttttg	catacatagc	acctctat	aatgcagaca	3360
gacttgctag	gctgaattcc	ttatcattgt	cacacacaca	tgaacaaaaa	taaaatatgg	3420
caggagagatt	ttagaaaccc	tactgtgaca	gcagttgtcc	atgcagtagt	cctgtaacaa	3480
ccagagacga	ggcattttcc	ctgagtgtgc	ttctcatcca	atcacacata	tccattcatt	3540
ccttcaagca	acatttgcctg	agtgtatact	atgttctagg	ggtaatgagt	aaaaataatt	3600
gttctgaaat	tacaatcctt	gcagtatctt	ttttaaaaga	gtgggtgaat	tttattctct	3660
gtgtacaata	gtaggagtga	aatatgggtac	tttcttattc	cagctaacgc	tattatttaa	3720

atgtatcatg aatcttttga gcaaataatg cataggatta ttaaagctat tattaataac	3780
acatgttaat tgttataatt atgatgttat attgtggagt tttcaatagt tctttggcac	3840
atgttgagaa gtagaaaata gctatgactt gattattaaa ttatttgcaa acagctgtag	3900
tgaattattht cgttaccaca tacaattcct caatgctgaa atgtgggatg aaaacataca	3960
cattgggttg agttttattht atagtattgc ttaataatag actgacacaa acattcatta	4020
cagaaaaatc tcaattctta ctaaggtagc aatgatttag ctggtctgcc aacaatatca	4080
gtgtagtcat tttttcaaaa ctaacatata aaaaagttha aagatcacat ttgatacatc	4140
ttgagtaatt atatctthaa agttataaat aaactgaagc atgtgaaaaa catgtcttga	4200
tttactgtag atgaatattc gggatgagaa aaatgaggac ttttgctcat cagatgagaa	4260
acctagaagc ctgtagccct agatthttacc tggagtthctg ccaaagggtta tattacaaa	4320
gccaatccaa tccatctcta cttcaacttht ttaatttgca aaacaaagat aaaaagtaag	4380
attaatcaac caattthaat ttaataagaa aatgcattta aggacatttht tatattthcaa	4440
ttgcaagata ctattcattg tattcatatt taatatagta aaatacattg gtgtthttatc	4500
acagcagatt gtaagtthaa tattatagtht ctcttgccaa thtcatththa aatgcatttht	4560
aattthtaaac agcaattthgt gggtcacaag ctgtcatcag attaaaattth atcacagtta	4620
tctaataaca ttgttagggc thttcatcag aactaaaaca aacacagctc cttggtctaa	4680
gcttctatac atttgagat taagaaaata agtggatggg atgacaaaga tcaagaaatc	4740
aagatattaa tctctacttg ttgaaggtgt gagthtgtha cthttatgga agactthtct	4800
ctcagaacag aggtgactt cataaagaag ggaaatgaat tctatacatt caggcctctt	4860
taactgtaaa gctaattgat ccattthtgt tacgcccagg gtatcagcaa ctgctthtatg	4920
ctggtggaac tgatgacacc ttataatgct ttagagatgg gaggaatgtg ctaacgggat	4980
gtaagtgtc tthtttagcta	5000

<210> 4

<211> 5000

<212> ADN

<213> Homo sapiens

<400> 4

gacgaccctg aatgatccct gcctcctagt acttatactc ttgtgtagcc tccacctact	60
tagactcaca gctgatttht tgaccaagag tataatggtg tatgactthca gaggctthgt	120
tcagagtcag aagtggccat thcatcatcc tctaaattthc tctthggcag attaatgaat	180
gtgagaaagt aaaaatggta tagaagaaag ttagcctgag agatggctgg cthtgtacct	240
gtthggtctt thtctthcct thcctccata agcaaatact gacgcatgta catatagaca	300
cactcaacac agtgggatgt taaagggaac ctagagccac tcgcccctthc ccaagatcag	360
agaacatagg aaagatgatg actthtttgca tctacctct cctctthgaac tcgcaacact	420
gtctthgaac acaaataagg tcaagtagga taaatctaca actgtctcct caactgtccg	480
ctcacaagtg tgaaattcaa ccattcctthg tgaagthgtc thttaatatc agtaatcagc	540
aataaccaac ctctgagctt tgcttatatt attaaatcat thtattthgac cctcaggaca	600
aagctgcaag cccagcatta cattctthcct thtatggaca aaaaaaaaaa aacaaaacaa	660
aaccgagata ggcacgaaga aactaataaa tggctgaaat gctattthgaa tccacatctg	720
tctaattcta aaagtgtgct thccaccaca tcaccgcacc tacaaggaaa atgcagcctc	780
taccctctcc tgaagaatgt gtaaaggagc aatatggtgt gthttggaaa aagcatcttht	840

gcagaaagaa	agcattagct	tcagtcatat	tagcatcctt	atccagcagc	agtagcatta	900
ttttacatat	cagttactcc	tccccaacaa	ctcaatgaat	tagatgccac	aatcatcctc	960
actttcaa	gtggagaccg	aggcacatct	ctaagactgg	agttcaggca	ttctggctcc	1020
agtcttagag	atgggttaagg	gttcacactc	ttaaccattt	attacaccat	agagctcacc	1080
aggtttgagg	gaaacaggat	caaatacaaaa	gagtcactca	ggactccagt	cctcactcaa	1140
ggacaaactg	ttccacctcg	gacagggaga	gtttccgcat	tctgagaccc	agcataacag	1200
gtcctgaccg	gcatctggca	ctcggactcc	caatcatact	ggatcacact	ggctcgggat	1260
gtgtaaagtc	caggggttct	cacatttgat	gacaccaaag	ccgcctaaaa	acaagagaga	1320
attaacaact	acctacggcg	gtctgatatt	tgcccaagag	atgccgcccc	ataaaactcc	1380
tttacatctt	tataacgttt	ttattttgcg	ttctccttca	taaccacat	ttaactcacc	1440
atagatgtaa	tgtttaaaat	tagttaccag	ataaactctt	acgtttccaa	actttaaggt	1500
tccttcgaaa	ccttctggta	aaactgttgt	tccacggaaa	tgggaacgta	acggatgagg	1560
caatcttcca	cagccgcaca	cagttgtgta	tccaccgcta	aacgggtcca	gtcatacatt	1620
caacgacca	cgcggagtca	gaagctacca	ccacacactg	tcaaaatcac	gcacacacag	1680
tgacggcccc	ttgcccactc	ggtcactcgc	ccacaatctc	tcgctagaga	atcacacgca	1740
gatagcacac	ccagcaccac	agaccccagg	aagcaacca	gggactcgaa	cacacgaaca	1800
gcactcctcc	gcgcaactgc	caggcacgcc	tgcgctccggc	tcacctgaa	acatcgcgag	1860
atccggttc	aaggccgggc	tgctgccttt	acgcctaaag	actatgtttc	ccggaagaca	1920
ctgcggcgcc	ggccctatca	tggcgcagca	tcggtgtgct	ttgtgcgtct	gcgccatctt	1980
ccggctgcgc	acggcgaatc	caccggtacc	gtggtggaag	cgcgccctgg	gctgccgggg	2040
gcgcggccgc	ggtggcactt	ggacccgagg	aggcggcagg	tgagaggttc	cggagctttc	2100
caggcgtct	ggggtccagc	aggagctgg	gcccggggccg	ggtgggtctc	aggcctgaga	2160
agaacgcaga	cgtctcgcct	catcgtcgt	ctgtggcttt	accggcgtga	gactacattt	2220
cccgcgggcc	ctcgcggcgt	gcgctttctg	gcgccccctt	tctgcttcca	gcctcatagc	2280
ccaggtgcat	ggacccctta	ggtgggtgcg	caggggtctc	cgacccctg	aaattgcgga	2340
tgttttgcat	tcaactgttat	gcgtgccttt	tttttttttt	aatctgagag	gaaatcttgt	2400
ttcatgaggt	tctcagagag	ttacaagacc	ccagaagact	tagaaccgct	agtttagaaa	2460
aacttatact	tgggaacatt	ttatatgttc	aaattctatg	tccagcaatg	gggaaacact	2520
tgagtggatc	acaggccatc	ctttgtaaat	aggatatcgt	gcagttaata	caaataatgc	2580
ttataaaagt	ggtcataata	aggaaaagct	gcctattaca	tcatattaag	caacattcat	2640
aattatttca	cttacacgtt	acttagaacc	aaatgtttgt	aatgcagtag	gcacettata	2700
tactccttta	cagagtactt	cttgtttaac	agtgagctca	agaactgggt	aaatgacaac	2760
actgcgaata	cttgggtgtt	ttttcttcca	atTTTTTTgt	tttgttttgt	tttttagcctg	2820
tgtgttagtg	agtttgactt	ctttttgctg	aatgttctca	aaggattcag	tagttaatat	2880
ctttgcttgt	cagtttttga	tagtcatttt	tgatagtcac	ttggcatccc	attatatggc	2940
tgtgccaaaa	atgattcaac	ctcttggtca	cttggtggact	tgccacttta	cgtactaatt	3000
ttcagttctc	ctatttagca	aatattattg	gacatgcagc	tttgtgcact	gctgtgtctc	3060
tatggaagga	cacatttatg	tgctgataat	ggagcactgg	gcgtgctgtg	tgtaatttgt	3120
ctcttactgt	tccttccaac	tctatcccat	ttgttgatgc	aaggcatgct	ccatgaagtg	3180
gcatgcacaa	agcctcagtt	tgggctgtag	agctttttgt	atatttggcg	gttgggaaaa	3240
gagccaggta	gactttcaaa	ggtgtttatat	ttgctgcagg	agagagttaa	atggcctagt	3300
cctggaaagg	ctttattcta	ggctggggct	atgcaaatga	tgcaggcctt	ctatgtctcc	3360
tggcttgagg	gaagccctcc	ttgtccttga	ggtaacatgt	gggtctcagt	cactagaaca	3420
gccctcagca	tccctgtctt	agacttgcat	ttgcagcaaa	acagaaccac	tggatgttta	3480

ggggaaatgg gagggaaagg accactgggt cccctttgcc ttgcagcagc taaagctggt 3540
 ttcttgggggt tgggtgagga tccctttgcc tccccctcct tctcattttc ttggtgggca 3600
 gagggcaagc aggtggggac ttgtaacttg gcattctagg aactgtgagg caagcctagc 3660
 ctctcatgtt tgctctttcc tccccagct ctaccgtcgc aggactctgc ctttccccaa 3720
 gagaggaacc agaaggagga cctatcagcc cttgaaattc taaaagtcac gtcccttgtg 3780
 agtttttaaa tcccataaat atttgctttc cttatcttga aactttccac ctgggttcca 3840
 tccagaaatt tggttctcag actttcccat tttagggaat aaaaggacca ggcaggagtc 3900
 tcctctccca ttgtcctctc ccctctgaca aatgcccaaca gatttattct ctatagcatt 3960
 ctgcactctc tgatggtttt cgtctgctag gaggagaaac tcatttccct gtatgaaatg 4020
 agagcatcct ttgtaagaac tgaagtttga agagatggaa tgaatgctta ggagcaatca 4080
 catatagaaa agcaggggag gtggtgttgg ttggtttcag tttcagcgca tggtagctg 4140
 tcctgctttg ggaatgatta atgattggat atgtaattga ggtttggctg ggaatccatc 4200
 aaaaatttga tgccagaatt aaagtctagt ttgacgttgg acagagcatt tgaatctggg 4260
 acatcctcca ttagtttctt tagcttaaga ggaagaaaag tacagagctg aggatggtga 4320
 gcgatgtagc aagtgagatg atcatctagg tgggtgagtg gtcagggcct ggaaaagcct 4380
 tagagatctg atccaacctc gggatccaaa ataagaggat gtttgagtca gttgtttcac 4440
 ttgatagaaa attgtgcaat caataaaagt tatctccatg gtttttccca tgaagaaaaa 4500
 tgcttgtgtt atgttccaac ataattgatga actttcatct ccaagaaatt cttgacttac 4560
 atgaaaatcc tgtatctaag atagaaaaca tgtttccata attttaaaaa ataaaatatt 4620
 tggggaaaat ccatcagaaa tattttttgt tgttattgtt attccaaaag atagaagtga 4680
 taggtctcac cttgagaatt tattgtatgc caagtataga gtgggcagtg tgctttatta 4740
 cactgtttct taacaataag gaaaactcca tttagcttga ttttaatttt catcccagac 4800
 ctgtacacag agcattgatt tgcagttaaa aggaatgttt gagaacaatt tgatcattct 4860
 gttttactca tccccatttc ttctgtcacc tttcacattc agtcccaccc ttcttgttca 4920
 acaaaacaac cccccctccc gcaaagacct gccccatctc ctttcatccc actgtgaacc 4980
 attgaaatac atatatatca 5000

<210> 5

<211> 27

<212> ADN

<213> trình tự nhân tạo

<220>

<223> mỗi xuôi

<400> 5

gttttcgtcg tttttcgttc ggagatc

27

<210> 6

<211> 25

<212> ADN

<213> trình tự nhân tạo

<220>

<223> mỗi ngược

<400> 6

gctctcgctc tcgctattac tcgct

25

<210> 7

<211> 23

<212> ADN

<213> trình tự nhân tạo

<220>

<223> mỗi xuôi

<400> 7

tgtgtatgcg cgttttttcg ttc

23

<210> 8

<211> 23

<212> ADN

<213> trình tự nhân tạo

<220>

<223> mỗi ngược

<400> 8

acctatatat ccgccgctcc gtc

23

<210> 9

<211> 24

<212> ADN

<213> trình tự nhân tạo

<220>

<223> mỗi xuôi

<400> 9

cggcggttggg tatgttttagt agtc

24

<210> 10

<211> 29

<212> ADN

<213> trình tự nhân tạo

<220>

<223> mỗi ngược

<400> 10

aattacgaat aaaaaaaaca aaaacgctc

29

<210> 11

<211> 28

<212> ADN

<213> trình tự nhân tạo

<220>

<223> mỗi xuôi

<400> 11

tgacggtttt ttgtttattc ggttattc

28

<210> 12

<211> 21

<212> ADN

<213> trình tự nhân tạo

<220>

<223> mỗi ngược

<400> 12

cgaacgcaaa cgtacctacg c

21

<210> 13

<211> 26

<212> ADN

<213> trình tự nhân tạo

<220>

<223> mỗi xuôi

<400> 13

ggttaagttt ttttttyggy gtagtt

26

<210> 14

<211> 26

<212> ADN

<213> trình tự nhân tạo

<220>

<223> mỗi ngược

<400> 14

tactccctct acctcccrac tctctc

26

<210> 15

<211> 28

<212> ADN

<213> trình tự nhân tạo

<220>

<223> mỗi xuôi

<400> 15

ttgtggygta gaggattatt agtttggt

28

<210> 16

<211> 22

<212> ADN

<213> trình tự nhân tạo

<220>

<223> mỗi ngược

<400> 16

ctaaaaacrc aaccatccc tc

22

<210> 17

<211> 27

<212> ADN

<213> trình tự nhân tạo

<220>

<223> mỗi xuôi

<400> 17

ggaattttat ttgaaattt ttttggt

27

<210> 18

<211> 33

<212> ADN

<213> trình tự nhân tạo

<220>

<223> mỗi ngược

<400> 18

ccccacttca aataaaatac tattaaaaaa aac

33

<210> 19

<211> 30

<212> ADN

<213> trình tự nhân tạo

<220>

<223> mỗi xuôi

<400> 19

tagtgayggt tttttgttta ttyggttatt

30

<210> 20

<211> 26

<212> ADN

<213> trình tự nhân tạo

<220>

<223> mỗi ngược

<400> 20

taaacrtaaa aacaacaacc cracct

26

<210> 21

<211> 5000

<212> ADN

<213> không biết

<220>

<223> giải trình tự bisulfit của DBC1

<400> 21

ggtttttatt tgaagttggt tgattaygtt gtttttagatt gatttttagaa ttgtattatt	60
agtagggtag ttgttgttta ttaatagtta tggaagtaga ttttagttgg gggatgatgg	120
ggaaataaag agaaagatag ggaygaataa aggaaggaat ggggtataga tagaaagaga	180
gagagagaaa agaaagaaa gaggagaggt aagagatagt gatttgggggt tttatatygy	240
ggtgggtggg agaaggagtg tgtgtgggtg gygtattttt tygtgggttag aattaaagtt	300
agaattgatt tggggaatgt gtgtttaaga tttgaatttg gattagaagt taagggaag	360
agaaagttgt ttaagaaaaa agggagaaat ttttattagg ygaaaataga aatygttgat	420
tgggtttgtg gttggagagt gttygtggt gttgattttt ttttattggg attttagtga	480

aggttgagaa	gttttgtatt	ggttgttttt	tttttttttt	tttttttttt	ttagttttga	540
ttttttgaga	tttaggatta	ttttggttag	ttygatgttt	tttttttttt	tttttttttt	600
ttttttgtyg	attaaataat	ttttttygag	tttttagttt	attattttaa	tatttttttt	660
ttattttgtt	ttaagatttt	ttttttttta	attagtttat	tggtttttag	attgttttag	720
gtttttttgg	tggttatatt	ttttttttat	tattttttga	attaaaaaat	aaataaaaaa	780
tatataatgg	taattataaa	agtagataat	atgtattgag	tatttattag	gtgttaggga	840
gtagtttaag	tgtttttgta	tatattaatt	tatgagtagt	ttaagtgttt	ttgtatatat	900
taatttatta	taattttata	ttaattttat	tagatgggta	gatatattat	tattttatta	960
aatagataag	gaattgaggt	ttaaggggta	agtgatttta	ttaagattag	atagttatag	1020
gattttaatt	tttttggagt	ttgtattttt	attgaatttg	gtttttttat	atattttaat	1080
attttttttt	atgatttttt	gtttttttat	ttggtagttt	taatgttttt	tatttttggt	1140
taatttttag	gggattaggg	atttgtattg	ttgatttggt	ttaaattaga	agggaagtaa	1200
agtaaaataaa	tatattaaat	gtgygttggt	tgggaattat	tataygaggg	ttttattttt	1260
tgtttagga	agaggtttta	tgtagtagt	tttagtttgt	atttaggttg	attgtagagt	1320
attttgtttt	ttatttttat	gttttagttt	ttgtattttt	gttttttttt	ygtttttggt	1380
ggtttttaga	gaatttygtg	tttttttagt	tttttttttt	tatattttgt	ttaygtagag	1440
aagtttttgt	tttatttttg	gaggttaygt	gggttttygt	ttatatatyg	agagaaataa	1500
atagtgttaa	atatttatag	agagaygygt	agatataaay	ggatttatay	gggtaatttt	1560
ygagataaaa	tttatattyg	atggatttay	gyggytggtg	aaatatttgt	ygtttttagaa	1620
atatttaggt	attygygata	tatatagtat	agttaygttt	aagggtatta	ggatttyggg	1680
tttygygta	tgygygggtt	ttttggatgt	tygtgygtat	agatataata	ttttataygt	1740
tttagattta	ygaaattttt	tayggtttag	tttttagttta	ttygggygtg	tttttttytg	1800
aggyggtttt	tygttttttt	ttttttygtt	tttttttttt	tttygtttta	agatgtataa	1860
aatattttty	ggaagtaatt	tyggygttta	gttttttttt	tttygttttt	yggtygtygt	1920
ttttttattt	atttttyggy	gtygtygggt	aagttttttt	ttyggygtag	ttyggttttt	1980
gtygtttttt	gttyggagat	ygyggygtat	ttggattttt	ttttttatty	gttagtygtt	2040
tygttttygg	atttttayggt	tgtaaattga	tttgggygyg	ggggaggagg	agagygtagg	2100
ygagygaatt	ygygagagag	ggagagagyg	agygagtaat	agygagagyg	agagygagag	2160
agtygggagg	tagagggagt	agtgatyggt	tttyggagty	gggatttatg	tttgttttyg	2220
ggattagyga	aggggatttt	ayggttgagt	atgagttagg	ttgttaggag	ttaggtattt	2280
ttaygtttgt	agtttttygy	tygtgttygg	aatgygagtt	gtaygtaggg	tttttttaag	2340
tttttatyga	gtygaataaa	aagygttttt	tygtagtttt	tygttaaaga	yggatattga	2400
tttttaggtaa	ggyggygtyg	ggtgtagygt	ttygtagttt	ygttgttttt	ggatttygggt	2460
tygggtygta	ttyggggygt	tttygygttt	ttttgttttt	ttttttatyg	gtatttttgt	2520
tygtttttta	tttggytggt	tygtyggttt	taagtttttt	ttagtttttag	ggaggggggt	2580
tttgtgtttg	gggttttaaag	ggttaattgy	gggttttagt	gagtgggygtg	tgtgttttyg	2640
ygygttttyg	aygtgtgtat	tatggtggga	attgatgtgt	gtgttagtgt	gtttgygtgt	2700
gyggygtygt	ttgygtttga	tgygtgtgtt	ygtggtgtgt	gtgtgtgtty	gtgtgtgtaa	2760
gggaggggtg	aagagagaga	ggttttataa	tttatttayg	gygygatgtg	tgtgtgtatg	2820
tttgtaygta	tgtgtttgta	tgtgtgtttt	ygtgtgtttt	tttaattagg	tttttttagt	2880
ttatayggaa	tgggattttt	attataggat	taygtagtta	tygggaaatt	ygttgtggat	2940
ttttttttgg	ggttttgggt	ttggggtttg	gggaggatta	tggggttgta	gatgggtatt	3000
tatttagttt	aatgttggtg	ygtttgaagg	aaaatttttt	taaayggtgg	aattttgtta	3060
tattgggatt	tatagtttaa	tatgaaagag	atatggttaa	tttytgaggt	aaatgagayg	3120

```

ttgttat ttt aaat tttata ttgggtagat ttttaagattt tgatgggaat ttggagatat 3180
tggatagtgg gtgatagaga aagggggaag ttagygggtgg gttttttatt tggggtttgg 3240
aaagtttggg atagggattt attttgtatt ttygtttgta attaatagag ttttttygg 3300
ttttgttgg tttggggagg atttggttaag tttgtatgga ttttttttag gggaaaggag 3360
aaagygtttt yggggattt tttattttatt atagtgtgtaa aggggttttag aggaaatttt 3420
atttggggyg gttgtggatg atatggaagg agatggatgg ggtatttttt taagatagat 3480
tygttttttt tttttttatt taggygggga agtttttaga gagtttttt ttaaataatgt 3540
ttttttatgg tttttttgga gttaggggga tattgagaga aggtagagag gtggagaggg 3600
agagagagag agtaagagy agagagaggg agatagagag agagagtgtt ttagtattga 3660
gggagattta taatttgaag aggggttgtg agtgtggaat ttatgataga atgtggtagt 3720
aattgattta aattgttgy ggtttgtatt tygygttttt tatttggttt taaaatattg 3780
ttaagggtag gggggyggg gaggaatttt agttaggaag gtagtttggg ttaaggtttt 3840
tttttttttt tttttttatt tttgggagg ttttttaggta ggttttttgg ttatttttgg 3900
ttggtattgt tgggttttggg ttttgggggt agttattgat ttgattttta gttgtttttt 3960
aaaggtttgt tttatttagt aaaaatgttg agttttattg ttgttagtat ggttttttaga 4020
ttyggtagt gttatttttt taaggtaagy ggtagtygtt ttgttttttag gtagggagg 4080
taggagaaga taagggttgt gaygttggga tatagttttt tgttttagttt agtayggagg 4140
ggygttgaa aagttttttt aggggaaaaa agtagaaata tatttggttt aaaaatgta 4200
tatatatatt ttaggaagat atatatatat atatatatat taattatata tatatatata 4260
tatatatatg tgtttatatt ttggggagag ataaaagtaa atatgtattt gtaggaattt 4320
gttttagtgg gtgggtttta tttgtaggta tttgtagata tattattttt tgttttagagg 4380
ttgatatttt aggttgaagt ttttatatag ttatagagaa ttttttttag gaaatttatg 4440
ttagggtaga gttttttgga ttatttttgg taggggggtat tagggttaag gaggagattt 4500
atttttggga attgttttat tttttttttt ttttaattta tagaygatyg aaattggaag 4560
gaatattaaa gagtagaaag gtaggatttt gtggttttaa gagggaatga gaatttttta 4620
atatttagta gtgggttagt ggtatgttta ggattagaat ttaggtttat taattagtta 4680
gattaatatt tttttttatt ttttttagtaa tatatatatt aggattgtat attgttttta 4740
tttttttttt gtttaattgt tattatgatt tttattaaga taatygtttg ttttttttta 4800
ttaaatttta tattttttaga gatgtatttt agggtttttt tatttggaaa tatggttttg 4860
aagagaaaag atttaagtgt ttaattatta ttttygaaag atatatatat atatatatat 4920
atatatatat atatatatat atatatatat atatataygg tgggtgggggg aggattgaga 4980
gagagataga gagagaatat 5000

```

<210> 22

<211> 5000

<212> ADN

<213> không biết

<220>

<223> giải trình tự bisulfit của PDE8B

<400> 22

```

agaataattt tgaagtatta tttttttttt agaatttttt ttgtagggat gtttgagggt 60
tggttaattt ttttttttgt ttaagtttgt ttttttttat tttttttttt tatttttttt 120

```

tataaatgtt	gattataaga	gtatTTTTta	aaaagtTTTT	TTTTgtttgt	taaatattat	180
TTtagaattt	gttttttagga	ggatttaatt	tgtgtagaaa	TTTTatgagt	aatttgtttg	240
TTtattaaaa	aatatagtat	agattTTTTt	taggttagtt	ggagatggag	TTTTtttgta	300
TTaaaaaaag	agagttaaat	tgagaagaga	agtaaaggta	gaagaataag	gtagattttt	360
TTTTgtataa	ttgaatttgg	gattgggaat	gatattttgt	TTTTttatga	aatttagaaa	420
ttgtTTTTtg	aagTTTTtat	ttagttatat	ttaagagtgg	tttaggaaga	aaaaggaaaa	480
taagaatgaa	aaaaaaataa	aaaataaatt	ttaaaattgg	atatTTTTtt	gtgttttgta	540
ttaagttttt	atgttattgt	TTTTgttygt	ataagtaatt	ttagataaaa	TTTTTTTTta	600
gattatttaa	ttattTTTTt	ygtgttattg	tgaaaatygg	gtaggaatta	ttgygggaat	660
tagtgattga	atgggtagtg	gtgtttygta	gttgtaattt	agagtttata	agggaaggat	720
gaatagtagt	tattTTTTta	aaggtatttt	TTTTTTTTtt	tagtatagtt	tagtttgtgt	780
atgttttatt	aaggagaatg	agggaaggtt	ttgttgattg	tttagtttag	agtgtgagga	840
atagtagttg	atatgggygg	ttggayggga	tagaaattag	tatgaatatt	tagtgatatt	900
taaggaagag	gtgaaaatgg	gggagaaagt	taagaattat	tagatagttt	TTTTttgttt	960
ttgttttgga	TTTTagggga	aygtatgtgt	ttaagggtgt	TTTTtttygt	agggtTTTTt	1020
atttttttgt	ataaattttt	aaaygatgag	tagatatgga	atgggtTTTT	tttgaagatt	1080
gggttattat	ttaatttagt	tttyggagta	ggyggtgggg	ttgggggaag	aggayyggaat	1140
tatgaggttt	TTTTattatt	tttygggtgt	gttgagaggt	aagtaaatgy	ggtgggtggg	1200
TTTgggttgt	agggtTTTTg	ttatttyggg	agggtTTTTt	tggtgttttag	TTTTgggatg	1260
tgtgaaaatg	tgttggtgaa	aagygagagg	tttatatagt	TTTTtagggg	aagaggggtt	1320
gggygttgg	gggyggygty	gggatgagtg	tagaagagay	gaggTTTTtg	gatagyggag	1380
gaggagggga	gggygtygag	gygyggtgtt	agttgtygyg	tataggggtt	tygyggygga	1440
gtygagtygy	gggtaygttt	tgTTTTgtyg	ggagagtTTY	gggagygggy	ggaggggygg	1500
agggygtag	tggggttygg	gyggttygyg	tygyggagty	gggtattttg	aggaggaagg	1560
agggtgggag	ygagggaggg	aggggayggg	ygtagatyga	aagtggggaa	agaagggtga	1620
ggtaggyggg	taggygggyg	ggygtTTTTg	tttagggtyg	ygggtgyggg	agttygggya	1680
ggytagttg	ggyggygggy	gggtygygt	ygagggagga	ggggaagggy	gaggygyggg	1740
gagygtgttt	gggygtygy	ggyggggagg	gtggyggyg	ttggtgygyg	ygggygttg	1800
tgtatgygyg	TTTTtygtt	yggggaggaa	gatggtttaa	aagggaaggt	tggggtgayg	1860
ygygyggttt	tyggaggtty	ggyggggggg	atygyggtta	gttygaygga	gyggyggata	1920
tataggttyg	gggygygta	gttygggygt	ygygyggty	gtTTTTttat	tgtaggtggt	1980
agygggtgyg	ttgggtTTYg	gyggttygyg	gygygggygg	gygygygggg	gagttyggtt	2040
gagggatggg	ttgygtTTTT	agtatttatg	tttygtagag	yggyltgatt	tattgtyggg	2100
attyggayga	gtttagttyg	tttygttaga	ttattagygt	gtygtagggg	tyggygggat	2160
TTTTgttygg	TTTTtygtt	tagatygayg	tygygaygt	tattTTTTtyg	agtygygygt	2220
ygggattttt	tagygtagtt	ygygttygta	gggttygtat	ygagttgggt	agyggtagta	2280
gygygggttt	ygtagtTTYt	gtygygatta	ttagtagggg	tyggaggygt	tattgttgta	2340
gtagygyga	ggytagatt	tagatttggt	atattagygt	gaaggtaaat	gtttygygtt	2400
ggtataygty	gtgggggtyg	ttygtTTYgt	ygygggggt	ygtaygggta	gggggtTTYg	2460
gyggagttag	gtgatygtga	ggyggttggt	ttggagaggt	tgttattaag	gaggagtTTa	2520
TTTTttattt	gtggagatga	tgggagTTta	ggaaatgtgg	ttagaaaaag	gtTTTTggag	2580
gggtTTTTga	agygtTTTTa	gttggtTTTT	ggggattggg	yggggaaggg	agygtagaag	2640
gaagtaggtg	ggttggtttg	TTTTTTTTtg	agggtaggaa	ggttggtggt	tggtttatgt	2700
aggaagaggg	gtggggatta	ttgagagtat	tyggtggtta	gttttggtga	atgaaatttg	2760

agtattgagt	tggatttgyg	tgttttgtag	gtgattggtg	tagttgtagt	attaggatag	2820
atagtgtttt	atatttygat	ttttatttgg	gattatttagt	taggttggag	ttttagtata	2880
ggaatygagy	gtagggattt	gtgaatgaat	gagtgttygt	gttttaagag	atgtgggaay	2940
ggagtagagt	ggaatttgtt	gtttgttatt	gtaaygtttt	tttgggttgg	ttgtatttta	3000
gatagaattg	agagaatygg	gttatgagtt	yggagtgtta	gtagagttat	ygtgagggga	3060
ygtggttttt	agtgtagtat	agttgtttga	ggatgatttt	gtatatataa	ttgatttttt	3120
tagagagtgg	gattttgagg	aagtggaaaa	aattgttttag	atggttaaag	tagtygttaa	3180
atatttatatt	tttaaagata	gaagaaaaat	aattatttta	atagtgtttt	ttygtatgtt	3240
tttaaagtat	ygttaaattt	aagaggtttt	ttattgtgta	aatttgggta	ttgggttttt	3300
ttttttttta	gtaaataaga	taataatggt	atattttatt	gtatagagag	aaaaaaatat	3360
ttttaatgtt	agatagaatt	atagttttta	tttggttaat	aggttattaa	gaattatttt	3420
atggattatt	tgtaaagat	atatttgtat	ttttaatagt	taaaaatttg	gtttttattg	3480
gatgaaggta	gtttatagtg	gaggagttag	gagaaggata	atatgttttg	gttttaattt	3540
ggaagtttta	aagtttttta	atgaatttat	ttttttttta	atagtatttg	attgtttaat	3600
atgaagtttg	attgatgttt	gtttgtggtt	gtagtgtttt	tttggttagta	attataatta	3660
ttattatagt	aaaaatatyg	tttattgagt	atattattatg	tgtaggagt	aagttttttt	3720
ttttaattat	tgaatgttaa	tagtgtgatt	ttagtatagt	atgtttgaag	attagagttt	3780
aggatttatg	atagtataaa	tttattagtt	gtttatttygt	tatagaggta	tgtgttgaay	3840
gttatatttt	gtttttgtta	agtttttttt	tttttttttt	ttttgtttat	ataaagttag	3900
ttttaggagt	aatttttagg	gatgaaaatg	ttaaatttgg	tgaaaatatt	aagttgggat	3960
attattagta	aagattattt	ttttgtttga	yggttgaaga	tttgaagtag	tagtttaaaa	4020
tttgtatagt	ttatgaattt	tgtgatattt	tttttggttg	ttttagagaa	attgtgaatt	4080
tttgtttttt	agaatattta	taatttttagt	ttagtatttg	gtaatttttag	ttattttatg	4140
tttgtttagt	agaagaaaga	atattgaatg	aattttatttt	gtatatatttt	aatatttttt	4200
ttaagaggtt	gttttagagat	taaatgaagt	aaaaattgaa	gggataatga	gttttaatta	4260
ttagtagttt	gtataaaattt	gtattaattt	tattgttatt	tggtttaagt	attaaaaggt	4320
ttgatttggg	taggagttta	atattatttg	ttttattaag	aagattttta	aataagagtt	4380
taatttatatt	taatttatta	ttgggaatgt	tggaatgaa	atttttgggt	tattgattta	4440
atattatttt	gtatttgttg	gatttttagaa	agtatatatt	tgaagtgaaa	tttgagatta	4500
gagttattta	tgtaaatttt	atagaattat	tattatagaa	ttgttttaat	atatttaata	4560
gtattttttt	aggtttgtga	ttygggtgtg	tattgttttt	ttttaatygt	aaagtttata	4620
tgttttgtgt	ttttttgttg	ttgagataat	tttataaaaa	ttttatatatt	atttttttgt	4680
agattaagta	tttttttttt	gtygtaagag	tttgagatag	ttgttaggta	ygtgggtggtg	4740
ttttattttt	ttttaagta	tataggtgga	aaaataaaat	ttttattatt	tggttttgtt	4800
tatagtgggtg	gaaggttgta	atagttttta	tttttttttt	ttggttattt	tatattaatt	4860
atgaagtgtg	tttttattat	ttatttgggt	tttggtattt	tttagagtgt	aaaaatgtgt	4920
gtttgatttt	tttttttttt	tttttttagta	ataattttatt	ttattaagtt	tttagaatga	4980
tgaggtatta	tgatgttaaa					5000

<210> 23

<211> 5000

<212> ADN

<213> không biết

<220>

<223> giải trình tự bisulfit của PTPRR

<400> 23

aagaaaaata gatttattgt aagtgggata taattatttt tttttaattt ttagttttgt	60
agtattttta gtaatagatg ttagatgaga agagtatttt ttttaagttt tatagaatgt	120
taaataaatt attttaaagg ttgatttagt aagaatttta aagagttaat tgtagtatt	180
gaattttgtt aattattatt tttatttatt tatttttgag tatttatttt gtatgaggta	240
ttgttttagg gagttgggat atattaataa agttattaaa gttttttaat attggagttt	300
aaattttgtt ttgygggggg gagtagggaa atagatatta aataatagat tttataagta	360
attattttgg tttgttagaa agtgataygg gaaaagtaaa ttggaggata gtgaattagg	420
aataggagaa gaggtttagt tttagagtgg tttgggatga aatgatagtg ttatagaaat	480
ggagtgtgtt ataygtatgg gggagttagg ggtggagtay gttgttagag aggggttaga	540
tatggagata tgtttttatt ttttagtagag aaggtttgtt aagatattaa aagtataatt	600
tattgttttt tttttttatt tgttggagtg ttttttggtt ttgtttttta gagttgaaga	660
tgatataaat tataaaagaa ataaatggta tygtttgggt tttttgattt tataaggttt	720
tatttaattt aagttttggt ggatagttta ttttttatta gttgttttaa tttatagagg	780
atggagaaaag aaagtattga atgagatgtg tagtttaatt taatgggtat ttgttagttt	840
taattatttg gttaggataa ttttatttaa gttaatggaa tataaaattt agtagttttt	900
atggagattt taatgtgttg tagaataaat tatagataaa tgaattgttt gaaagaataa	960
tatgaggtag ggtaaataa gtgttggtga agagaaagat tattgtaatt tagagttagt	1020
tagagttagg gttttttgga aaagtगत taatattata ttttygaaaa ttagtgtggt	1080
ggtggtgtaa gtggaggaaa atgggaaaga ggtaatagtg tttataaaga tatagaggta	1140
gaaagggtag aggttttggg atagttaaaa gagaagtta gttgagggat tttttgattt	1200
gattagtagt aattagttat gattaataag gttttttata ttttaaagat ttagatggag	1260
ggataaaaaa ttatttaatg gtagatttg gtagtttttt tagagatata gagttgggty	1320
ggatgagttt aggtattgay gtgatttatt attttttatt ttaaagagta aaagggaat	1380
taaagttaat tatttttayg aaataaaaag gtgttttttt gtgttttaat tatatggata	1440
tattttatta gtttaaaaagt atttttttat tttttttgt tattgtgagg atttgagtta	1500
gaagaaagt taaatatagt tattgagttg gaaagagtgg aaagagaagt aaagaggggg	1560
aagtttagg aaggaygaag ttatttttaa gatatatggt tattgtttat attaagtaag	1620
ttgttttggg aaygtttttt tygagttagt agaatttta gtagtggaag atattttttat	1680
ttttgtaggy gagttttggg aagttggtta atttgtaaat gttatttttt agtagtgagt	1740
tyggtttayg tgtaaaataa gatttgggga aagagttagg tgggtggtat ggttgataat	1800
gttattagtt tttttttttg atttttgtgg tygtgttttt atttattttt atttagttat	1860
attttatttt yggatttggt atggaygttg ggtttttagt aattatagta agtgtttttt	1920
tygtattttt ttttttttta tttttatttt tatttttaat tattatttta gygatggagt	1980
ttattttgtt ttaagtygty gttaagatty ggagaagygg aattttattt tgaaattttt	2040
ttgttttggt agggtyggyg ttgggtatgt ttagtagtyg ygygttggt gttgggtgt	2100
tgggttggyg yggagtttat tttgtygttt tygttttggt ttttgggygt ttagaagggt	2160
aggtatttgt ygtttttgag ygtttttgtt ttttttatty gtaatttttt attgtttttt	2220
tttttttttt tttttaggga ggttgaagtt ggtgttggtt tttgtyggyg ttatagattg	2280
attgttttgt aaatttttagt ygaggatttg aatttyggag attagaagat ttttggyggt	2340
ggtttttttt aatagtattt tatttgaagt ggggtygtgg tggagttttt tttttatttt	2400

ttaatgtaaa	tattatgygg	agagtagttt	gtttttttgy	gttgtgtttg	ttttttaatt	2460
tttaygttgt	aggtaagggg	ttgttaggtt	tagagtygga	gttgtgtatg	agatgggaaa	2520
ttgtatatgt	ttaaggattt	taggaaaagt	ttgttttgyg	aagaaaagttt	ttttaaaaaag	2580
gtaaaatagg	atagtatttg	atagggaagg	ggtgaggggag	tytgttattt	gttggaaaat	2640
tgaggtygaa	aatttaattt	aaaattagtt	tttgattttt	ttttatttta	taygaaaata	2700
gtgaatattt	ttaatatgaa	tataggtttg	tttttttttt	ygtttttttt	tygttttagat	2760
tttttgtttt	tgtttttagg	ttatgtattt	aattgttaat	tttttaggag	agggataaga	2820
tattttgtta	gatgtaattt	tttttattgt	agyggttatt	atattttata	ggtttttttg	2880
tttagygagy	gtttatagga	aggtattggt	tgtaatttga	tggattatat	tttygtttaa	2940
aagatygaaa	tatatggttt	atatttagtga	tgttgtatta	ggtttttttg	gtttttgttt	3000
ttgttattyg	aattttttta	gttaataaga	ggaagaaaag	gttttaaaata	gaattttttt	3060
gtttttttta	gtatygtggt	tagygttygt	tggtatatta	aaaaaaaaaa	aaaaaaaaaa	3120
aaaaaaaaaa	ataattatag	ttttaaaagt	gataattttg	ggtttaggtat	tgtttttygg	3180
atagatttga	taaggtgtaa	ttaagagtaa	tagtgagatt	gtttttaaat	ataaatagat	3240
ttgaggtatt	agaattaata	ttaagtttyg	tgaggaaagta	aattttggga	agtatttgag	3300
agatattata	tggtgtgttt	ttgttttttg	tatatatagt	attttttattt	aatgtagata	3360
gatttgtag	gttgaatttt	ttattattgt	tatatatata	tgaattaaaa	taaaatatgg	3420
tagggagatt	ttagaaattt	tattgtgata	gtagttgttt	atgtagtagt	tttgtaataa	3480
ttagagayga	ggtatttttt	ttgagtgtgt	tttttattta	attatatata	tttatttatt	3540
tttttaagta	atatttggtg	agtgtatatt	atgttttagg	ggtaatgagt	aaaaataatt	3600
gttttgaaat	tataattttt	gtagtatttt	ttttaaaaaga	gtgggtgaat	tttatttttt	3660
gtgtataata	gtaggagtga	aatatgggtat	ttttttattt	tagttaaygt	tattatttta	3720
atgtattatg	aattttttga	gtaaataatg	tataggatta	ttaaagttat	tattaaatat	3780
atatgttaat	tggtataaatt	atgatgttat	attgtggagt	ttttaatagt	tttttggtat	3840
attttgagaa	gtagaaaata	gttatgattt	gattattaaa	ttatttgtaa	atagttgtag	3900
tgaattattt	yggtattata	tataattttt	taatgttgaa	atgtgggatg	aaaatatata	3960
tattgggttg	agttttattt	atagtattgt	ttaataatag	attgatataa	atattttatta	4020
tagaaaaatt	ttaattttta	ttaaggtagt	aatgatttag	ttggtttggt	aataatatta	4080
gtgtagttat	ttttttaaaa	ttaatatata	aaaaagttta	aagattatat	ttgatatat	4140
ttgagtaatt	atatttttaa	agttataaat	aaattgaagt	atgtgaaaaa	tatgttttga	4200
tttattgtag	atgaatatty	gggatgagaa	aaatgaggat	ttttgtttat	tagatgagaa	4260
atttagaagt	ttgtagtttt	agattttatt	tggagttttg	ttaaagggtta	tattattaaa	4320
gttaatttta	tttattttta	ttttaatttt	ttaatttgta	aaataaagat	aaaaagtaag	4380
attaattaat	taattttta	ttaataagaa	aatgtattta	aggatatatt	tatattttaa	4440
ttgtaagata	ttattttattg	tatttatatt	taatataagta	aaatatattg	gtgttttatt	4500
atagtagatt	gtaagtttaa	tattatagtt	tttttgttaa	ttttatttta	aatgtatttt	4560
aattttaaat	agtaatttgt	gggttataag	ttgttattag	attaaaaatt	attatagtta	4620
tttaataata	ttgttagggt	tttttattag	aattaaaata	aatatagttt	tttggtttta	4680
gtttttatat	atttgagat	taagaaaata	agtggatggg	atgataaaga	ttaagaaatt	4740
aagatattaa	tttttatttg	ttgaagggtg	gagtttgta	tttttatgga	agattttttt	4800
tttagaatag	aggttgattt	tataaagaag	ggaaatgaat	tttatatatt	taggtttttt	4860
taattgtaaa	gttaattgat	ttattttggt	taygtttagg	gtattagtaa	ttgttttatg	4920
ttggtggaat	tgatgatatt	ttataatgtt	ttagagatgg	gaggaatgtg	ttaaygggat	4980
gtaagtgttt	tttttagtta					5000

<210> 24

<211> 5000

<212> ADN

<213> không biết

<220>

<223> giải trình tự bisulfit của ZNF582

<400> 24

gaygattttg aatgattttt gtttttttagt atttatattt ttgtgtagtt tttattttatt	60
tagattttata gttgattttg tgattaagag tataatgggtg tatgatttta gaggttttgg	120
ttagagttag aagtgggttat tttattattt tttaaatttt tttttggtag attaatgaat	180
gtgagaaagt aaaaatggta tagaagaaag ttagtttgag agatggttgg ttttgtattt	240
gtttggtttt tttttttttt tttttttata agtaaatatt gaygtatgta tatatagata	300
tatttaatat agtgggatgt taaagggaat ttagagttag tygttttttt ttaagattag	360
agaatatagg aaagatgatg attttttgta tttatttttt ttttttgaat tygtaatat	420
gttttgaatt ataaataagg ttaagtagga taaatttata attgtttttt taattgttyg	480
tttataagtg tgaaatttaa ttatttttgt tgaagttgtt ttttaatat agtaattagt	540
aataattaat ttttgagttt tgtttatatt attaaattat tttatttgat ttttaggata	600
aagttgtaag tttagtatta tttttttttt tttatggata aaaaaaaaaa aataaaaataa	660
aatyagata ggtaygaaga aattaataaa tgggttgaaat gttatttgaa tttatatattg	720
tttaatttta aaagtgtgtt ttttattata ttatygtatt tataaggaaa atgtagtttt	780
tatttttttt tgaagaatgt gtaaagaggt aatatgggtg gttttggaaa aagtattttt	840
gtagaaagaa agtatttagtt ttagttatat tagtattttt atttagtagt agtagtatta	900
ttttatatat tagttatttt tttttaataa tttaatgaat tagatgttat aattattttt	960
atttttaaat gtggagatyg aggtatattt ttaagattgg agtttaggta ttttggtttt	1020
agtttttagag atggttaagg gtttatattt ttaattattt attatattat agagtttatt	1080
aggtttgagg gaaataggat taaattaaaa gagttattta ggatttttagt ttttatttaa	1140
ggataaattg ttttatttyg gatagggaga gttttygtat tttgagattt agtataatag	1200
gttttgatyg gtatttggtt ttyggatttt taattatatt ggattatatt ggtyyggat	1260
gtgtaaagtt tagggttttt tatatttgat gatattaaag tygttttaaaa ataagagaga	1320
attaataatt atttayggyg gtttgatatt tgtttaagag atgtygtttt ataaaaattt	1380
tttatatttt tataaygttt ttatttttyg ttttttttta taatttatat ttaatttatt	1440
atagatgtaa tgtttaaaat tagttattag ataaattttt aygtttttta attttaaggt	1500
tttttygaaa ttttttggtt aaattgttgt tttayggaaa tgggaaygta ayggatgagg	1560
taatttttta tagtygtata tagttgtgta tttatygtta aayggtttta gttatatatt	1620
taaygattta ygyggagttt gaagttatta ttatatattg ttaaaaattay gtatatatag	1680
tgayggtttt ttgtttatty ggttattygt ttataatttt tygttagaga attataygta	1740
gatagtatat ttagtattat agatttttagg aagtaattta gggattygaa tataygaata	1800
gtattttttt gygtattgyg taggtaygtt tgygttyggt ttattttgaa atatygygag	1860
attyggtttt aaggttygggt tgttgttttt aygttttaag attatgtttt tyggaagata	1920
ttgyggygty ggttttatta tgygtagta tyggtgtgtt ttgtgygttt gygttatttt	1980
tyggttgygt ayggygaatt tatyggaty gtggtggaag ygygttttgg gttgtygggg	2040

gygyggtygy ggtggtat	ggattygagg aggyggtagg	tgagaggttt yggagttttt	2100
taggygtttt ggggtttagt	aggagttagt gtttgggttg	ggttgggtttt aggtttgaga	2160
agaaygtaga ygttttgyt	tatygtgtt ttgtggtttt	atygygtga gattatattt	2220
ttygttgggtt ttygyggyt	gygtttttttt gygttttttt	tttgttttta gttttatagt	2280
ttaggtgtat ggatttttta	ggtgggtgyg taggggtttt	ygattttttt aaattgygga	2340
tgttttgtat ttattgttat	gygtgtttttt tttttttttt	aatgtgagag gaaattttgt	2400
tttatgaggt ttttagagag	ttataagatt ttagaagatt	tagaatygtt agtttagaaa	2460
aatttatatt tgggaatatt	ttatatgttt aaattttatg	tttagtaatg gggaaatatt	2520
tgagtggatt ataggttatt	ttttgtaaat aggatatygt	gtagttaata taaataatgt	2580
ttataaaagt ggttataata	aggaaaagt gtttattata	ttatattaag taatatttat	2640
aattatttta tttataygtt	atttagaatt aaatgtttgt	aatgtagtag gtattttata	2700
tattttttta tagagtattt	tttgtttaat agtgagtta	agaattgggtt aaatgataat	2760
attgygaata tttgggtgtt	ttttttttta atttttttgt	tttgttttgt ttttagtttg	2820
tgtgttagtg agtttgattt	ttttttgtt aatgttttta	aaggatttag tagttaatat	2880
ttttgtttgt tagtttttga	tagttatttt tgatagtatt	ttggtatttt attatatggt	2940
tgtgttaaaa atgatttaat	tttttggtta tttgtggatt	tgtttattta ygtattaatt	3000
tttagttttt ttatttagta	aatattattg gatagttagt	tttgtgtatt gttgtgtttt	3060
tatggaagga tatatttatg	tggtgataat ggagtattgg	gygtgttgtg tgtaatttgt	3120
tttttattgt tttttttaat	tttattttat ttgttgatgt	aaggatatgtt ttatgaagtg	3180
gtatgtataa agtttttagt	tgggtttagt agttttttgt	atatttgyg gttgggaaaa	3240
gagttaggta gatttttaaa	ggtgttatat ttgtttagg	agagagttaa atggtttagt	3300
tttggaagg ttttatttta	ggttgggtt atgtaaatga	tgtaggtttt ttatgttttt	3360
tggtttgagg gaagtttttt	ttgtttttga ggtaatatgt	gggttttagt tattagaata	3420
gttttttagta tttttgtttt	agatttgtat ttgtagtaa	atagaattat tggatgttta	3480
ggggaaatgg gagggaaaagg	attattgggt tttttttgtt	ttgtagtagt taaagtgggt	3540
tttttgggt tgggtgagga	ttttttgtt tttttttttt	ttttattttt ttggtgggtg	3600
gagggtagt aggtggggat	ttgtaatttg gtatttttagg	aattgtgagg taagttagt	3660
tttttatgtt tgtttttttt	tttttttagt ttatygtgt	aggattttgt ttttttttaa	3720
gagaggaatt agaaggagga	tttatttagt ttgaaattt	taaaagttaa gttttttgtg	3780
agtttttaaa ttttataaat	atttgttttt ttatttttga	aattttttat ttgggtttta	3840
tttagaaatt tggtttttag	atttttttat tttagggaat	aaaaggatta ggtaggagtt	3900
ttttttttta ttgttttttt	ttttttgata aatgtttata	gatttatttt ttatagtatt	3960
ttgtattttt tgatggtttt	ygtttgttag gaggagaaat	ttattttttt gtatgaaatg	4020
agagtatttt ttgtaagaat	tgaagtttga agagatggaa	tgaatgttta ggagtaatta	4080
tatatagaaa agtaggggag	gtggtgttgg ttggtttttag	tttttagygt tggatatagt	4140
ttttgttttg ggaatgatta	atgattggat atgtaattga	ggtttgggtt ggaatttatt	4200
aaaaatttga tgttagaatt	aaagttagt ttgaygttgg	atagagtatt tgaatttggg	4260
atatttttta ttagtttttt	tagtttaaga ggaagaaaag	tatagagttg aggatggtga	4320
gygatgtagt aagtgagatg	attatttagg tgggtgagtg	gtaggggtt ggaaaagt	4380
tagagatttg atttaattty	gggatttaaa ataagaggat	gtttgagtta gttgttttat	4440
ttgatagaaa atttgttatt	taataaaagt tatttttatg	gtttttttta tgaagaaaaa	4500
tgtttgtgtt atgttttaat	ataatgatga atttttattt	ttaagaaatt tttgatttat	4560
atgaaaaatt tgtatttaag	atagaaaaata tgtttttata	attttaaaaa ataaaaatatt	4620
tggggaaaaat ttattagaaa	tattttttgt tgttattgtt	attttaaaag atagaagtga	4680

taggttttat	tttgagaatt	tattgtatgt	taagtataga	gtgggtagtg	tgttttatta	4740
tattgttttt	taataataag	gaaaatttta	tttagtttga	ttttaatttt	tatttttagat	4800
ttgtatatag	agtattgatt	tgtagttaaa	aggaatgttt	gagaataatt	tgattatttt	4860
gttttattta	tttttatttt	ttttgttatt	ttttatatatt	agttttattt	tttttgttta	4920
ataaaataat	ttttttttty	gtaaagattt	gttttatattt	tttttatttt	attgtgaatt	4980
attgaaatat	atatatatta					5000