



(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)



1-0042546

(51)⁸ C12Q 1/68; C12N 15/11

(13) B

(21) 1-2019-00467

(22) 29/06/2017

(86) PCT/CN2017/090740 29/06/2017

(87) WO 2018/001295 04/01/2018

(30) 201610509983.0 30/06/2016 CN

(45) 27/01/2025 442

(43) 25/04/2019 373A

(73) 1. CAPITALBIO CORPORATION (CN)

18, Life Science Parkway, Changping District, Beijing 102206, P.R.China

2. BEIJING CAPITALBIO MEDLAB CO., LTD. (CN)

Building D1, D2 Block 88, Kechuangliujie St, Beijing Economic-Technological
Development Area Beijing 101111, P.R.China

(72) GUO, Hongyan (CN); SUN, Yimin (CN); WANG, Yahui (CN); XIE, Zhan (CN);
XING, Wanli (CN); CHENG, Jing (CN); DENG, Tao (CN); ZHANG, Zhiwei (CN).

(74) Công ty Luật TNHH WINCO (WINCO LAW FIRM)

(54) BỘ ĐOẠN MÔI VÀ BỘ ĐOẠN DÒ ĐỂ KHUẾCH ĐẠI TỔ HỢP GEN VÀ GEN
THAM CHIẾU NỘI, KIT THỬ NGHIỆM CHỨA CHỨNG

(21) 1-2019-00467

(57) Sáng chế đề cập đến bộ đoạn môi và bộ đoạn dò để khuếch đại tổ hợp gen và gen tham chiếu nội, kit thử nghiệm chứa chúng. Bằng cách sử dụng thông tin theo dõi đối với phương pháp so sánh, độ chính xác của kit thử nghiệm theo sáng chế trong việc dự đoán nguy cơ ở các bệnh nhân được chẩn đoán ban đầu là mắc bệnh ung thư vú dương tính với thụ thể ER hoặc PR về sự tái phát hoặc tử vong trong khoảng thời gian từ 3 đến 10 năm sau khi phẫu thuật bằng 70%, và độ chính xác trong việc dự đoán nhóm có nguy cơ thấp và nhóm có nguy cơ cao tương ứng là 81,1% và 54,4%. Độ chính xác tương ứng trong việc dự đoán kết quả thử nghiệm bệnh lý FFPE tương ứng là 71,9% và 56,8%. Mô hình dự đoán nguy cơ dựa trên kit thử nghiệm này hỗ trợ chỉ cần giá trị chu kỳ ngưỡng Ct của gen đánh dấu phân tử, độ tuổi của bệnh nhân, giai đoạn pT, và số lượng LN, và không cần dựa trên các thông tin bệnh lý lâm sàng khác; mô hình này cung cấp sự đánh giá tiên lượng bệnh ung thư tốt hơn so với kết quả dự đoán bệnh lý một mình, và làm giảm sự xuất hiện việc điều trị không thích hợp ở mức độ nhất định do sự dự đoán bệnh lý sai gây ra, do đó cải thiện thêm phương pháp kỹ thuật để đánh giá tiên lượng bệnh ung thư.

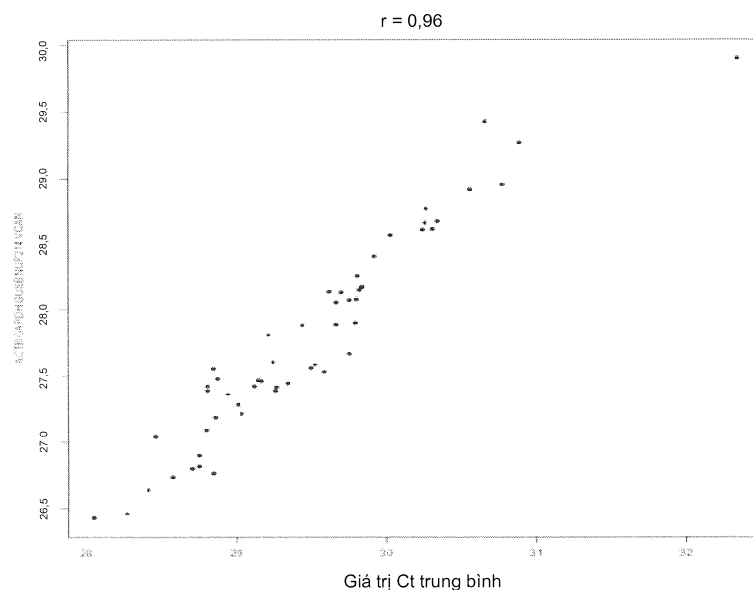


Fig.1

Lĩnh vực kỹ thuật được đề cập

Sáng chế liên quan đến lĩnh vực công nghệ sinh học, cụ thể là đề cập đến các gen đánh dấu phân tử, bộ đoạn môi và bộ đoạn dò để khuếch đại tổ hợp gen, kit thử nghiệm chứa chúng và phương pháp xây dựng mô hình thử nghiệm.

Tình trạng kỹ thuật của sáng chế

Bệnh ung thư vú là một trong những nguyên nhân chính đe dọa sức khỏe của phụ nữ trên khắp thế giới. Năm 2013, số liệu thống kê quốc gia về bệnh ung thư được công bố bởi Hiệp hội ung thư Mỹ cho thấy rằng tỷ lệ mắc bệnh ung thư vú đứng thứ nhất trong số các bệnh ung thư ở phụ nữ, và tỷ lệ tử vong đứng thứ hai. Theo số liệu mới nhất của Trung tâm ung thư quốc gia, trong năm 2013, có 232340 ca mắc mới bệnh ung thư vú được thông báo và 39620 ca tử vong được thông báo ở phụ nữ Mỹ. Ở Mỹ, trung bình cứ 8 phụ nữ có 1 người mắc bệnh ung thư vú. Mặc dù Trung Quốc là nước có tỷ lệ mắc bệnh ung thư vú thấp, tỷ lệ mắc và tỷ lệ tử vong do bệnh này đã tăng lên đáng kể trong những năm gần đây. Khoảng 15% trong số 1,3 triệu bệnh nhân ung thư vú mới được chẩn đoán mỗi năm trên thế giới là người Trung Quốc. Theo số liệu thống kê của mạng lưới ung thư vú ở Trung Quốc, tỷ lệ tăng hàng năm của bệnh ung thư vú ở Trung Quốc là từ 3% đến 4%, tỷ lệ này cao hơn từ 1% đến 2% tỷ lệ của thế giới. Tỷ lệ mắc bệnh này đứng thứ nhất trong số các bệnh ung thư dễ mắc phải ở phụ nữ. Cần phát triển nhanh chóng phương pháp phòng ngừa, chẩn đoán, tiên lượng, và điều trị được cá nhân hóa đối với bệnh ung thư vú.

Bệnh ung thư vú là một loại khối u không đồng nhất ở mức cao với nhiều yếu tố tiên lượng. Các bệnh nhân ung thư vú có cùng giai đoạn lâm sàng, phân độ mô học và mức độ biểu hiện của thụ thể hormon và được điều trị bằng phác đồ điều trị giống nhau có thể có sự tiên lượng khác nhau. Cách để xác định chính xác sự tiên lượng của các bệnh nhân ung thư vú và các sản phẩm tương ứng với chương trình điều trị được cá nhân hóa tương ứng để tránh sự ảnh hưởng có hại và gánh nặng do việc điều trị quá mức và điều trị không thích hợp gây ra là vấn đề cấp bách cần được giải quyết trong thực hành lâm sàng.

Với sự phát triển nhanh chóng của công nghệ sinh học phân tử, có thể tìm ra và phát hiện được các gen liên quan đến sự tiên lượng ở bệnh ung thư vú bằng cách sử dụng các phương pháp sinh học phân tử như phương pháp phản ứng chuỗi polymeraza (polymerase chain reaction: PCR), phương pháp lai đoạn dò và chip gen. Năm 2002, Van't Veer và các đồng tác giả đã sàng lọc 117 ca bệnh ung thư vú bằng công nghệ chip ADN và phát hiện được 70 gen liên quan đến sự tiên lượng bệnh ung thư vú. Năm 2004, các nhà nghiên cứu Mỹ đã sử dụng phương pháp phản ứng RT-PCR để kiểm tra 675 mẫu bệnh ung thư vú, và thu được 21 gen liên quan đến sự tiên lượng. Công ty Genomic Health đã phát triển Oncotype DX[®], sản phẩm liên quan đến tiên lượng bệnh ung thư vú. Oncotype DX[®] là sản phẩm xét nghiệm duy nhất để tiên lượng bệnh ung thư vú được khuyến cáo bởi hướng dẫn từ mạng lưới ung thư toàn diện quốc gia (NCCN), hướng dẫn lâm sàng của Hiệp hội ung thư lâm sàng Mỹ (ASCO) và sự đồng thuận lâm sàng của St Gallen. Ngoài ra, Yasuto Naoi và các đồng tác giả đã sử dụng công nghệ chip ADN để nghiên cứu các mẫu mô ung thư từ các bệnh nhân mắc bệnh ung thư vú dương tính với thụ thể ER và âm tính với hạch bạch huyết ở người Nhật Bản, và phát hiện được 95 gen liên quan đến sự tiên lượng. Nhóm nghiên cứu của Torsten O. Nielsen đã phát hiện được 50 tổ hợp gen có thể cung cấp thêm thông tin đối với sự tiên lượng dự đoán bệnh ung thư vú so với các yếu tố lâm sàng và kết quả nhuộm hóa mô miễn dịch, và sử dụng các mẫu FFPE thay vì mẫu tươi hoặc mẫu đông lạnh để thử nghiệm có thể mở rộng phạm vi của các mẫu có thể phát hiện. Các sản phẩm thử nghiệm liên quan đến sự tiên lượng bệnh ung thư vú-Oncotype DX[®], Mammaprint, ProSigna[™], và MapQuant Dx[™]-lần lượt được chấp thuận bởi Cơ quan quản lý thực phẩm và dược phẩm Mỹ (FDA) và Cộng đồng châu Âu (CE) trong các năm từ 2002 đến 2013. Tuy nhiên, các sản phẩm này hiện dựa trên sự nghiên cứu và phát triển đối với người châu Âu và người Mỹ. Khi đến Trung Quốc, các sản phẩm này trở nên đắt tiền, và không biết các gen và mô hình thử nghiệm của chúng có thích hợp để sử dụng cho người Trung Quốc hay không. Do đó, rất cần phát triển công nghệ thử nghiệm sự tiên lượng hữu hiệu và kinh tế đối với các bệnh nhân ung thư vú ở Trung Quốc.

Bản chất kỹ thuật của sáng chế

Để khắc phục các vấn đề nêu trên, sáng chế đề cập đến gen đánh dấu phân tử và sử dụng gen này, kit thử nghiệm, và phương pháp xây dựng mô hình thử nghiệm. Kit này

có tính năng tốt hơn trong thử nghiệm đánh giá việc chẩn đoán bệnh ung thư vú so với các kết quả đánh giá bệnh lý lâm sàng. Nó có thể làm giảm việc điều trị quá mức và điều trị không thích hợp do việc chẩn đoán bệnh lý sai gây ra ở mức độ nhất định, đáp ứng các yêu cầu của việc điều trị chính xác và được cá nhân hóa của các bệnh nhân ung thư vú, và cải thiện thêm các phương pháp kỹ thuật dự đoán tiên lượng của bệnh ung thư vú cho các bệnh nhân ở Trung Quốc.

Để đạt được mục đích nêu trên, sáng chế đề xuất giải pháp kỹ thuật sau đây.

Sáng chế đề xuất tổ hợp gen chứa (các) gen đánh dấu phân tử *MAPT* và/hoặc *MS4A1*.

Sáng chế đề xuất tổ hợp gen chứa các gen đánh dấu phân tử *BCL2*, *PGR*, *SCUBE2*, *ESR1*, *MKI67*, *CCNB1*, *MYBL2*, *GRB7*, *ERBB2*, *MMP11*, *CD68*, *BAG1*, *MAPT* và *MS4A1*.

Theo một số phương án của sáng chế, tổ hợp gen còn chứa các gen tham chiếu nội *ACTB*, *GAPDH*, *GUSB*, *NUP214* và *VCAN*.

Sáng chế còn mô tả việc sử dụng tổ hợp gen trong việc sản xuất dụng cụ thử nghiệm để dự đoán nguy cơ tái phát và/hoặc tử vong trong thời gian từ 3 đến 10 năm sau khi phẫu thuật ung thư vú.

Theo một số phương án của sáng chế, việc đánh giá và thử nghiệm nguy cơ tái phát và/hoặc tử vong trong thời gian từ 3 đến 10 năm sau khi phẫu thuật ung thư vú bao gồm các bước cụ thể là: thu được ARN tổng số của mẫu thử nghiệm, thu được ADN bổ trợ bằng cách phiên mã ngược, thu được giá trị chu kỳ ngưỡng Ct của gen đánh dấu phân tử và các gen tham chiếu nội bằng phản ứng PCR định lượng huỳnh quang, tính giá trị Ct trung bình của các gen tham chiếu nội để thu được giá trị Ct trung bình (Ct') của tổ hợp gen tham chiếu nội, chuẩn hóa giá trị Ct của các gen đánh dấu phân tử tương ứng bằng cách trừ đi giá trị Ct' của tổ hợp gen tham chiếu nội để thu được giá trị ΔCt , và sau đó phân tích giá trị ΔCt , và độ tuổi, giá trị pT và giá trị LN của đối tượng bằng mô hình dự đoán nguy cơ tái phát hoặc tử vong trong thời gian từ 3 đến 10 năm sau khi phẫu thuật ung thư vú được xây dựng bằng thuật toán rừng ngẫu nhiên để thu được kết quả. Trong số chúng, giá trị pT để chỉ giai đoạn bệnh lý, và giá trị LN để chỉ số lượng hạch bạch huyết di căn. Giá trị thu được bằng phương pháp phân tích này được so sánh với giá trị ngưỡng để thu được kết quả, trong đó giá trị ngưỡng bằng 5. Giá trị thu được bằng phương pháp phân tích này ≥ 5 thể hiện sự tiên lượng tốt, và giá trị thu được bằng

phương pháp phân tích này < 5 thể hiện sự tiên lượng xấu.

Theo một số phương án của sáng chế, phương pháp xây dựng mô hình dự đoán nguy cơ tái phát hoặc tử vong trong thời gian từ 3 đến 10 năm sau khi phẫu thuật ung thư vú bao gồm các bước: xây dựng ma trận bằng cách sử dụng giá trị ΔC_t của gen đánh dấu phân tử của mẫu thử nghiệm, và độ tuổi, giá trị pT và giá trị LN của đối tượng, chọn ngẫu nhiên một nửa làm nhóm tập huấn và một nửa còn lại làm nhóm kiểm chứng, thiết lập mô hình dự đoán chứa 10000 cây quyết định bằng thuật toán rừng ngẫu nhiên, lấy mẫu ngẫu nhiên ≥ 1000 lần và thiết lập ≥ 1000 mô hình dự đoán, chọn ≥ 39 mô hình tối ưu có tỷ lệ phù hợp cao nhất với thông tin theo dõi từ ≥ 1000 mô hình dự đoán làm mô hình con của mô hình cuối, và thiết lập giá trị trung bình của ≥ 39 mô hình con làm giá trị dự đoán cuối của nguy cơ tiên lượng.

Thuật toán rừng ngẫu nhiên gồm nhiều cây quyết định. Cây quyết định được xây dựng bằng phương pháp xác định ngẫu nhiên hai lần đối với thuộc tính và mẫu, vì thế, nó còn được gọi là cây quyết định ngẫu nhiên. Trong thuật toán rừng ngẫu nhiên, không có sự tương quan giữa các cây quyết định. Khi các số liệu thử nghiệm nhập vào rừng ngẫu nhiên, nó được phân loại bởi mỗi cây quyết định. Cuối cùng, danh mục có sự phân loại nhiều nhất ở tất cả các cây quyết định là kết quả cuối, nghĩa là kết quả “lựa chọn” của cây quyết định. Nói theo cách khác, thuật toán rừng ngẫu nhiên là công cụ phân loại bao gồm nhiều cây quyết định, và danh mục đầu ra của nó được xác định bằng phương thức của danh mục đầu ra của cây riêng biệt. Theo sáng chế, các tác giả sáng chế đã tối ưu hóa thuật toán rừng ngẫu nhiên truyền thống. Tác giả lấy mẫu ngẫu nhiên 1000 lần và thiết lập 1000 mô hình. 39 mô hình tối ưu có cả độ chính xác và độ đặc hiệu cao hơn từ 1000 mô hình được chọn làm mô hình con của mô hình cuối, và giá trị trung bình của 39 mô hình con được thiết lập làm kết quả dự đoán cuối cùng.

Độ tuổi của các bệnh nhân mới được chẩn đoán mắc bệnh ung thư vú dương tính với thụ thể ER hoặc PR giai đoạn trung gian hoặc không được điều trị sớm, giai đoạn PT, số lượng hạch LN di căn, và giá trị C_t của 14 gen đánh dấu phân tử *BCL2*, *PGR*, *SCUBE2*, *ESR1*, *MKI67*, *CCNB1*, *MYBL2*, *GRB7*, *ERBB2*, *MMP11*, *CD68*, *BAG1*, *MAPT*, *MS4A1* và 5 gen giữ nhà *ACTB*, *GAPDH*, *GUSB*, *NUP214*, *VCAN* được nhập vào 39 mô hình dự đoán của phép phân tích để thu được kết quả phân tích dự đoán, nhờ đó thu được giá trị nguy cơ tái phát hoặc tử vong trong thời gian từ 3 đến 10 năm, giá trị này sẽ được xác định là sự tiên lượng tốt hoặc sự tiên lượng xấu dựa trên ngưỡng nguy

cơ (ngưỡng bằng 5).

Theo một số phương án của sáng chế, mẫu thử nghiệm trong sáng chế là mẫu được cố định bằng formalin và gắn trong parafin (Formalin-Fixed and Paraffin-Embedded: FFPE) từ bệnh nhân mới được chẩn đoán mắc bệnh ung thư vú dương tính với thụ thể ER hoặc PR giai đoạn trung gian hoặc không được điều trị sớm.

Sáng chế còn đề xuất bộ đoạn môi để khuếch đại tổ hợp gen, các trình tự của chúng được thể hiện bởi các trình tự từ SEQ ID NO: 1 đến SEQ ID NO: 28.

Sáng chế còn đề xuất bộ đoạn dò để khuếch đại tổ hợp gen, các trình tự của chúng được thể hiện bởi các trình tự từ SEQ ID NO: 29 đến SEQ ID NO: 42.

Sáng chế còn đề xuất bộ đoạn môi để khuếch đại các gen tham chiếu nội của tổ hợp gen, các trình tự của chúng được thể hiện bởi các trình tự từ SEQ ID NO: 43 đến SEQ ID NO: 47.

Sáng chế còn đề xuất bộ đoạn dò để khuếch đại các gen tham chiếu nội của tổ hợp gen, các trình tự của chúng được thể hiện bởi các trình tự từ SEQ ID NO: 48 đến SEQ ID NO: 52.

Sáng chế còn đề xuất kit thử nghiệm để dự đoán nguy cơ tái phát và/hoặc tử vong trong thời gian từ 3 đến 10 năm sau khi phẫu thuật ung thư vú, kit này chứa bộ đoạn môi và/hoặc bộ đoạn dò, và các chất phản ứng thường được sử dụng trong kit.

Sáng chế còn đề xuất phương pháp xây dựng mô hình dự đoán nguy cơ tái phát hoặc tử vong trong thời gian từ 3 đến 10 năm sau khi phẫu thuật ung thư vú, bao gồm các bước: xây dựng ma trận bằng cách sử dụng giá trị ΔC_t của gen đánh dấu phân tử của mẫu thử nghiệm, và độ tuổi, giá trị pT và giá trị LN của đối tượng, chọn ngẫu nhiên một nửa làm nhóm tập huấn và một nửa còn lại làm nhóm kiểm chứng, thiết lập mô hình dự đoán chứa 10000 cây quyết định bằng thuật toán rừng ngẫu nhiên, lấy mẫu ngẫu nhiên ≥ 1000 lần và thiết lập ≥ 1000 mô hình dự đoán, chọn ≥ 39 mô hình tối ưu có tỷ lệ phù hợp cao nhất với thông tin theo dõi từ ≥ 1000 mô hình dự đoán làm mô hình con của mô hình cuối, và thiết lập giá trị trung bình của ≥ 39 mô hình con làm giá trị dự đoán cuối của nguy cơ tiên lượng.

Sáng chế còn đề xuất phương pháp đánh giá và thử nghiệm nguy cơ tái phát hoặc tử vong trong thời gian từ 3 đến 10 năm sau khi phẫu thuật ung thư vú, bao gồm các bước: thu được ARN tổng số của mẫu thử nghiệm, thu được ADN bổ trợ bằng cách phiên mã ngược, thu được giá trị C_t của gen đánh dấu phân tử và các gen tham chiếu nội

bằng phản ứng PCR định lượng huỳnh quang, tính giá trị Ct trung bình của các gen tham chiếu nội để thu được giá trị Ct trung bình (Ct') của tổ hợp gen tham chiếu nội, chuẩn hóa giá trị Ct của gen đánh dấu phân tử tương ứng bằng cách trừ đi giá trị Ct' của tổ hợp gen tham chiếu nội để thu được giá trị ΔCt , và sau đó phân tích giá trị ΔCt , và độ tuổi, giá trị pT và giá trị LN của đối tượng bằng mô hình dự đoán nguy cơ tái phát hoặc tử vong trong thời gian từ 3 đến 10 năm sau khi phẫu thuật ung thư vú được xây dựng bằng thuật toán rừng ngẫu nhiên để thu được kết quả, nghĩa là thu được giá trị nguy cơ tái phát hoặc tử vong trong thời gian từ 3 đến 10 năm và sự dự đoán sự tiên lượng tốt hoặc tiên lượng xấu dựa trên ngưỡng nguy cơ (ngưỡng bằng 5).

Theo sáng chế, mẫu thử nghiệm là mẫu FFPE từ bệnh nhân mới được chẩn đoán mắc bệnh ung thư vú dương tính với thụ thể ER hoặc PR giai đoạn trung gian hoặc không được điều trị sớm.

Các giải pháp kỹ thuật được sử dụng trong sáng chế bao gồm (1) bằng cách nghiên cứu tài liệu và cơ sở dữ liệu, chọn 192 gen ứng viên liên quan đến bệnh ung thư vú (không chỉ giới hạn ở các gen liên quan đến sự tiên lượng bệnh ung thư vú, chứa các gen tham chiếu nội) và tùy chỉnh các chip biểu hiện gen TLDA (Applied Biosystems Co.); (2) thu thập có hệ thống các số liệu nhân khẩu học đầy đủ, các số liệu lâm sàng và số liệu theo dõi (sự tái phát và thời gian di căn, thời gian sống), và chọn các mẫu FFPE của bệnh nhân mới được chẩn đoán mắc bệnh ung thư vú dương tính với thụ thể ER hoặc PR giai đoạn trung gian hoặc không được điều trị sớm, bằng cách sử dụng các chip TLDA được tùy chỉnh để thử nghiệm 192 gen và sàng lọc các gen tham chiếu nội và gen đánh dấu phân tử liên quan đến sự tiên lượng bệnh ung thư vú; (3) kiểm tra gen đánh dấu phân tử ứng viên và gen tham chiếu trong các mẫu độc lập và xây dựng mô hình dự đoán nguy cơ tái phát hoặc tử vong trong thời gian từ 3 đến 10 năm sau khi phẫu thuật bằng thuật toán rừng ngẫu nhiên, và đánh giá tỷ lệ phù hợp giữa mô hình dự đoán và kết quả theo dõi; (4) tiến hành kiểm tra thêm bằng cách sử dụng các mẫu lâm sàng độc lập: các mẫu FFPE của 19 trường hợp bệnh nhân mới được chẩn đoán mắc bệnh ung thư vú dương tính với thụ thể ER hoặc PR giai đoạn trung gian hoặc giai đoạn sớm hoặc có số liệu theo dõi lâm sàng đã biết và đánh giá tỷ lệ phù hợp giữa mô hình dự đoán và kết quả theo dõi.

Sáng chế đề xuất hệ thống thử nghiệm gen để đánh giá dự đoán sự tái phát hoặc tử vong trong thời gian từ 3 đến 10 năm sau khi phẫu thuật ở bệnh nhân mới được chẩn

đoán mắc bệnh ung thư vú dương tính với thụ thể ER hoặc PR giai đoạn trung gian hoặc không được điều trị sớm. Bằng cách chiết ARN tổng số từ các mẫu mô bệnh ung thư vú được cố định bằng formalin và gắn trong parafin (FFPE), sau khi phiên mã ngược bằng các đoạn môi thường dùng, giá trị Ct của sự biểu hiện của 14 gen đánh dấu phân tử liên quan đến việc chẩn đoán bệnh ung thư vú và 5 gen tham chiếu nội được phát hiện bằng cách sử dụng phản ứng PCR. Giá trị Ct và độ tuổi, giá trị pT và số lượng hạch LN của đối tượng được đưa vào mô hình dự đoán nguy cơ tái phát hoặc tử vong trong thời gian từ 3 đến 10 năm sau khi phẫu thuật ở các bệnh nhân ung thư vú dương tính với thụ thể ER hoặc PR giai đoạn trung gian hoặc giai đoạn sớm được xây dựng bằng thuật toán rừng ngẫu nhiên, để đánh giá sự tiên lượng tốt hoặc tiên lượng xấu. So sánh với thông tin theo dõi, hệ thống này đạt được độ chính xác 70%. Ngoài trừ độ tuổi bệnh nhân, giai đoạn pT, và số lượng hạch LN, không cần phải dựa vào các thông tin bệnh lý lâm sàng khác.

Kit được đề xuất trong sáng chế có tỷ lệ dự đoán chính xác bằng 81,1% đối với các bệnh nhân ung thư vú mới được chẩn đoán có nguy cơ thấp về sự tái phát hoặc tử vong trong thời gian từ 3 đến 10 năm, trong khi độ chính xác của thử nghiệm dự đoán bệnh lý bằng 71,9%. Ngoài ra, kit này tạo ra độ nhạy của tỷ lệ chính xác dự đoán bằng 54,4% đối với các bệnh nhân ung thư vú mới được chẩn đoán có nguy cơ cao về sự tái phát hoặc tử vong trong thời gian từ 3 đến 10 năm, độ chính xác này gần bằng độ chính xác của thử nghiệm dự đoán bệnh lý tương ứng bằng 56,8%. So với thông tin theo dõi trong lâm sàng, kit này đạt được tỷ lệ thích hợp bằng 70%. Ngoài trừ độ tuổi bệnh nhân, giai đoạn pT, và số lượng LN, không cần phải dựa vào các thông tin bệnh lý lâm sàng khác.

Hệ thử nghiệm và kit có tính năng tốt hơn trong thử nghiệm đánh giá dự đoán bệnh ung thư vú so với các kết quả đánh giá bệnh lý lâm sàng. Có thể làm giảm việc điều trị quá mức và điều trị không thích hợp do việc chẩn đoán bệnh lý sai gây ra ở mức độ nhất định, đáp ứng các yêu cầu về việc điều trị chính xác và được cá nhân hóa của các bệnh nhân ung thư vú, và cải thiện thêm các phương pháp kỹ thuật để dự đoán việc chẩn đoán bệnh của các bệnh nhân ung thư vú ở Trung Quốc.

Mô tả vắn tắt các hình vẽ

Để minh họa rõ ràng hơn các phương án của sáng chế hoặc giải pháp kỹ thuật của các giải pháp đã biết, các hình vẽ được sử dụng trong các phương án hoặc phần mô tả các

giải pháp đã biết sẽ được mô tả vắn tắt dưới đây.

Fig.1 thể hiện kết quả phân tích sự tương quan giữa các gen tham chiếu nội và gen thử nghiệm;

Fig.2 thể hiện việc xây dựng mô hình dự đoán nguy cơ tái phát hoặc tử vong trong thời gian từ 3 đến 10 năm sau khi phẫu thuật ung thư vú.

Mô tả chi tiết sáng chế

Sáng chế đề cập đến các gen đánh dấu phân tử, gen tham chiếu nội và ứng dụng của chúng, kit thử nghiệm và phương pháp xây dựng mô hình thử nghiệm. Người có hiểu biết trung bình trong lĩnh vực này có thể biết được từ nội dung của phần mô tả và cải thiện một cách thích hợp các thông số của quy trình. Cần hiểu rằng tất cả các phương án và cải biến này đều là hiển nhiên với người có hiểu biết trung bình trong lĩnh vực này và được cho là nằm trong phạm vi của sáng chế. Phương pháp và ứng dụng trong sáng chế đã được mô tả trong các phương án được ưu tiên, và hiển nhiên là người có hiểu biết trung bình trong lĩnh vực này có thể tiến hành thay đổi, cải biến và kết hợp thích hợp phương pháp và ứng dụng được mô tả ở đây mà vẫn nằm trong phạm vi của sáng chế, và tiến hành và áp dụng kỹ thuật của sáng chế.

Các giải pháp kỹ thuật được sử dụng trong sáng chế bao gồm (1) bằng cách nghiên cứu tài liệu và cơ sở dữ liệu, chọn 192 gen ứng viên liên quan đến bệnh ung thư vú (không chỉ giới hạn ở các gen liên quan đến việc tiên lượng bệnh ung thư vú, chứa các gen tham chiếu nội) và tùy chỉnh các chip biểu hiện gen TLDA (Applied Biosystems Co.); (2) thu thập có hệ thống đầy đủ các số liệu nhân khẩu học, các số liệu lâm sàng và số liệu theo dõi (sự tái phát và thời gian di căn, thời gian sống), và chọn các mẫu FFPE của bệnh nhân mới được chẩn đoán mắc bệnh ung thư vú dương tính với thụ thể ER hoặc PR giai đoạn trung gian hoặc không được điều trị sớm, bằng cách sử dụng các chip TLDA được tùy chỉnh để thử nghiệm 192 gen và sàng lọc các gen tham chiếu nội và gen đánh dấu phân tử liên quan đến sự tiên lượng bệnh ung thư vú; (3) kiểm tra các gen đánh dấu phân tử ứng viên và gen tham chiếu trong các mẫu độc lập và xây dựng mô hình dự đoán nguy cơ tái phát hoặc tử vong trong thời gian từ 3 đến 10 năm sau khi phẫu thuật bằng thuật toán rừng ngẫu nhiên, và đánh giá tỷ lệ phù hợp giữa mô hình dự đoán và kết quả theo dõi; (4) tiến hành kiểm tra thêm bằng cách sử dụng các mẫu lâm sàng độc lập: các mẫu FFPE của 19 trường hợp bệnh nhân mới được chẩn đoán mắc bệnh ung thư vú

đương tính với thụ thể ER hoặc PR giai đoạn trung gian hoặc giai đoạn sớm có số liệu theo dõi trong lâm sàng đã biết và đánh giá tỷ lệ phù hợp giữa mô hình dự đoán và kết quả theo dõi.

1. Chọn các mẫu nghiên cứu

(1) các bệnh nhân mới được chẩn đoán mắc bệnh ung thư vú giai đoạn sớm không được điều trị hoặc giai đoạn trung gian;

(2) các bệnh nhân ung thư vú dương tính với thụ thể ER hoặc (và) PR giai đoạn sớm không được điều trị hoặc giai đoạn trung gian;

(3) có hoặc (và) không có sự di căn hạch bạch huyết, và số lượng hạch bạch huyết di căn;

(4) có thông tin theo dõi chi tiết và chính xác.

Tổng cộng có 339 mẫu đáp ứng các tiêu chí được sử dụng trong nghiên cứu này.

2. Chiết ARN tổng số của mẫu FFPE

ARN tổng số được chiết từ các mẫu FFPE bằng cách sử dụng kit phân lập ARN FFPET có độ tinh khiết cao (Roche), và ARN thu được có nồng độ nằm trong khoảng từ 25 đến 400 ng/ μ l, mật độ quang OD 260/280 nằm trong khoảng từ 1,8 đến 2,0, và OD 260/230 nằm trong khoảng từ 1,5-2,0.

3. Thử nghiệm chip TLDA (Applied Biosystems).

Các thử nghiệm sau đây được thực hiện đối với 26 cặp mẫu có sự tiên lượng tốt và sự tiên lượng xấu từ các thông tin theo dõi trong lâm sàng đã biết:

(1) tiến hành phản ứng phiên mã ngược đối với ARN tổng số để thu được mẫu ADN bổ trợ;

(2) tiến hành thử nghiệm chip TLDA đối với mẫu ADN bổ trợ;

(3) phân tích và xử lý số liệu để thu được các gen đánh dấu phân tử ứng viên và gen tham chiếu nội.

4. Tiến hành phản ứng RT-PCR định lượng trong thời gian thực (Real-time quantitative RT-PCR: qRT-PCR)

Các gen đánh dấu phân tử ứng viên được kiểm tra đối với 289 mẫu từ các thông tin theo dõi trong lâm sàng:

(1) tiến hành phản ứng phiên mã ngược đối với ARN tổng số để thu được mẫu ADN bổ trợ;

- (2) tiến hành thử nghiệm RT-PCR đối với mẫu ADN bổ trợ;
- (3) phân tích và xử lý số liệu.

5. Phương pháp chuẩn bị kit chẩn đoán

Bằng phương pháp tùy chỉnh chip TLDA, các gen tham chiếu nội và gen có mức độ biểu hiện khác nhau giữa 26 mẫu bệnh ung thư vú có sự tiên lượng tốt và 26 mẫu bệnh ung thư vú có sự tiên lượng xấu được xác định, và các gen tham chiếu nội và gen biểu hiện khác biệt được sàng lọc. Các gen đánh dấu phân tử ứng viên được kiểm tra với lượng lớn các mẫu bằng phản ứng PCR định lượng huỳnh quang phiên mã ngược. Cuối cùng, 14 gen và 5 gen tham chiếu nội (*BCL2*, *PGR*, *SCUBE2*, *ESR1*, *MKi67*, *CCNB1*, *MYBL2*, *GRB7*, *ERBB2*, *MMP11*, *CD68*, *BAG1*, *MAPT*, *MS4A1*; *ACTB*, *GAPDH*, *GUSB*, *NUP214*, *VCAN*) liên quan đến sự tiên lượng bệnh ung thư vú được sàng lọc, các gen này được sử dụng để tạo thành kit chẩn đoán. Kit chẩn đoán này chứa các đoạn môi và đoạn dò của các gen này, và các chất phản ứng thông thường khác dùng cho phản ứng qRT-PCR. Kit còn chứa mô hình dự đoán, trong đó mức độ biểu hiện của các gen *BCL2*, *PGR*, *SCUBE2*, *ESR1*, *MKi67*, *CCNB1*, *MYBL2*, *GRB7*, *ERBB2*, *MMP11*, *CD68*, *BAG1*, *MAPT* và *MS4A1* được thử nghiệm bằng cách sử dụng giá trị trung bình của các gen *ACTB*, *GAPDH*, *GUSB*, *NUP214* và *VCAN* làm gen tham chiếu. Đồng thời, các thông tin lâm sàng như độ tuổi của bệnh nhân ung thư vú, giá trị PT và LN được sử dụng để đánh giá tổng quát nguy cơ tái phát hoặc tử vong trong thời gian từ 3 đến 10 năm sau khi phẫu thuật ung thư vú, và sự tiên lượng tốt hoặc sự tiên lượng xấu được xác định.

6. Thiết lập mô hình đánh giá nguy cơ

- (1) Thiết lập mô hình để dự đoán nguy cơ tái phát hoặc tử vong trong thời gian từ 3 đến 10 năm sau khi phẫu thuật ung thư vú

Thuật toán rừng ngẫu nhiên trong phương pháp học bằng máy được sử dụng để đánh giá và thử nghiệm giá trị nguy cơ tái phát hoặc tử vong của các mẫu trong thời gian từ 3 đến 10 năm sau khi phẫu thuật, và thiết lập mô hình thử nghiệm gen dùng cho việc đánh giá chẩn đoán bệnh ung thư vú. Thuật toán rừng ngẫu nhiên gồm nhiều cây quyết định. Cây quyết định được xây dựng bằng phương pháp chọn ngẫu nhiên hai lần đối với của các thuộc tính và mẫu, vì thế nó còn được gọi là cây quyết định ngẫu nhiên. Trong thuật toán rừng ngẫu nhiên, không có sự tương quan giữa các cây quyết định. Khi số liệu thử nghiệm nhập vào rừng ngẫu nhiên, nó được phân loại bằng mỗi cây quyết định. Cuối cùng, danh mục có sự phân loại nhiều nhất trong tất cả các cây quyết định là kết quả cuối

cùng, nghĩa là kết quả “lựa chọn” của cây quyết định. Nói theo cách khác, thuật toán rừng ngẫu nhiên là công cụ phân loại bao gồm nhiều cây quyết định, và danh mục đầu ra của nó được xác định theo phương thức của danh mục đầu ra của cây riêng biệt. Theo sáng chế, các tác giả sáng chế đã tối ưu hóa thuật toán rừng ngẫu nhiên truyền thống. Tác giả sáng chế lấy mẫu ngẫu nhiên 1000 lần và thiết lập 1000 mô hình. 39 mô hình tối ưu có cả hiệu quả và độ đặc hiệu cao hơn từ 1000 mô hình được chọn làm mô hình con của mô hình cuối, và giá trị trung bình của 39 mô hình con được thiết lập làm kết quả dự đoán cuối cùng.

Trong pha thứ nhất của nghiên cứu, tổng cộng 192 gen ứng viên liên quan đến bệnh ung thư vú được thử nghiệm bằng kỹ thuật TLDA. Sự khác biệt về mức độ biểu hiện gen giữa 26 mẫu bệnh ung thư vú có sự tiên lượng tốt và 26 mẫu bệnh ung thư vú có sự tiên lượng xấu được phát hiện, và các gen có mức độ biểu hiện khác nhau được sàng lọc. Sự khác biệt về mức độ biểu hiện gen được thể hiện là $2^{-\Delta Ct}$, trong đó $\Delta Ct = CT_{\text{mẫu}} - CT_{\text{tham chiếu}}$, và các gen tham chiếu được sử dụng làm thông số tham chiếu để chuẩn hóa và tính toán mức độ biểu hiện tương đối. Quá trình sàng lọc gen tham chiếu chuẩn nội bao gồm các bước: sàng lọc gen tham chiếu nội ứng viên bằng cách sử dụng các chương trình phần mềm geNorm, BestKeeper, NormFinder, delta Ct dựa trên thuật toán ổn định và xem xét chức năng sinh học của các gen ít biến động và mối liên quan của chúng với khối u; tính toán sự tương quan giữa giá trị Ct trung bình của tất cả các tổ hợp gen tham chiếu chuẩn nội ứng viên và giá trị Ct trung bình của 192 gen, và tổ hợp có mối tương quan cao nhất là tổ hợp gen tham chiếu chuẩn nội, chứa các gen *ACTB*, *GAPDH*, *GUSB*, *NUP214* và *VCAN*. Các tiêu chuẩn sàng lọc của gen ứng viên bao gồm (1) phép phân tích tổng quan- mức độ khác biệt giữa hai nhóm có sự tiên lượng tốt và sự tiên lượng xấu là gấp 2 lần hoặc nhỏ hơn 0,5, và tỷ lệ các trường hợp có giá trị $Ct < 35$ bằng 50%; (2) phép phân tích phân tầng – liên quan đến sự không di căn hạch bạch huyết, mức độ khác biệt giữa hai nhóm có sự tiên lượng tốt và sự tiên lượng xấu là gấp hơn 2 lần, và sự khác biệt về mặt thống kê là $< 0,05$; (3) trong phép phân tích tổng quan, mức độ khác biệt giữa hai nhóm là không đáng kể, tuy nhiên điều này đã được thông báo trong tài liệu chuyên ngành liên quan đến sự tiên lượng bệnh ung thư vú, và tỷ lệ các trường hợp có giá trị $Ct < 35$ bằng 90%. Các gen đáp ứng các tiêu chuẩn sàng lọc trên đây bao gồm *BCL2*, *PGR*, *SCUBE2*, *ESR1*, *MKI67*, *CCNB1*, *MYBL2*, *GRB7*, *ERBB2*, *MMP11*, *CD68*, *BAG1*, *MAPT* và *MS4A1*.

Theo phương án thứ nhất, sáng chế đề xuất gen đánh giá sự chẩn đoán bệnh đối với các bệnh nhân ung thư vú ở Trung Quốc. Hiện nay, các sản phẩm tương tự ở nước ngoài được phát triển trên cơ sở người châu Âu và người Mỹ, và các nhóm dân tộc khác nhau có mức độ biểu hiện gen khác nhau. Theo sáng chế, 19 gen được sàng lọc, trong đó các gen *MAPT* và *MS4A1* liên quan đến việc dự đoán sự tái phát hoặc tử vong trong thời gian từ 3 đến 10 năm sau khi phẫu thuật dựa trên số liệu về các bệnh nhân ung thư vú ở Trung Quốc. Mặc dù các gen đã được thông báo là có liên quan đến bệnh ung thư vú, không có báo cáo trực tiếp về vai trò của nó trong việc tiên lượng bệnh ung thư vú. Theo phương án thứ hai, sáng chế thiết lập tổ hợp gen tham chiếu chuẩn nội mới, tổ hợp này khác với các ứng dụng và sản phẩm khác, và ít bị ảnh hưởng bởi chất lượng ARN trong các mẫu FFPE, làm cho kết quả thử nghiệm của các gen đánh dấu phân tử đáng tin cậy hơn. Theo phương án thứ ba, mô hình dự đoán trên cơ sở thuật toán rừng ngẫu nhiên được sử dụng cho phép phân tích toàn diện, tiến hành dự đoán nguy cơ tái phát hoặc tử vong trong thời gian từ 3 đến 10 năm sau khi phẫu thuật ở bệnh nhân mới được chẩn đoán mắc bệnh ung thư vú có ER⁺ hoặc PR⁺ giai đoạn sớm hoặc trung gian.

Do các vấn đề nêu trên, sáng chế đề xuất hệ thử nghiệm gen đánh giá tiên lượng đối với sự tái phát hoặc tử vong trong thời gian từ 3 đến 10 năm sau khi phẫu thuật ở các bệnh nhân ung thư vú dương tính với ER hoặc PR giai đoạn I và giai đoạn II không được điều trị. So với thông tin theo dõi, hệ thống này đạt được độ chính xác 70%. Ngoài trừ độ tuổi, giai đoạn pT, và số lượng LN của bệnh nhân, không cần phải dựa vào các thông tin bệnh lý lâm sàng khác.

Các nguyên liệu ban đầu và chất phản ứng được sử dụng trong sáng chế của các gen đánh dấu phân tử, gen tham chiếu nội và sử dụng chúng, kit thử nghiệm và phương pháp xây dựng mô hình thử nghiệm đều có bán trên thị trường.

Sáng chế được minh họa thêm dựa vào các ví dụ sau đây.

Ví dụ thực hiện sáng chế

Ví dụ 1: lấy mẫu và thiết lập thông tin mẫu

Các mẫu FFPE của các bệnh nhân ung thư vú mới được chẩn đoán không được điều trị được chọn, và số liệu theo dõi trong lâm sàng được thu thập có hệ thống đầy đủ. Bằng cách thiết lập thông tin mẫu, 341 mẫu đáp ứng các tiêu chuẩn sau đây được chọn làm mẫu thử nghiệm cho thử nghiệm chip kiểu mảng mật độ thấp Taqman (Taqman Low

Density Array: TLDA) và loạt các phản ứng qRT-PCR sau đó:

- (1) các bệnh nhân mới được chẩn đoán mắc bệnh ung thư vú ở giai đoạn trung gian hoặc giai đoạn sớm không được điều trị;
- (2) các bệnh nhân ung thư vú dương tính với thụ thể ER hoặc (và) PR giai đoạn trung gian hoặc giai đoạn sớm không được điều trị;
- (3) có hoặc (và) không có sự di căn hạch bạch huyết, và số lượng hạch bạch huyết di căn;
- (4) có thông tin theo dõi chi tiết và chính xác.

Ví dụ 2: sàng lọc các gen đánh dấu phân tử và gen tham chiếu nội bằng chip TLDA

Thử nghiệm chip TLDA được thực hiện đối với 26 mẫu bệnh ung thư vú có sự tiên lượng tốt và 26 mẫu bệnh ung thư vú có sự tiên lượng xấu đáp ứng tiêu chuẩn nêu trên, và thu được kết quả. Các bước cụ thể là như sau.

(1) Chiết ARN từ các mẫu FFPE: chuẩn bị 4 lát cắt có độ dày 20µm hoặc 8 lát cắt có độ dày 10µm cho mỗi mẫu; ARN được chiết theo quy trình của kit phân lập ARN của FFPE có độ tinh khiết cao (Roche); và ARN chiết được đạt yêu cầu về sự kiểm soát số lượng và chất lượng bằng cách sử dụng NanoDrop-2000 được cho thử nghiệm phiên mã ngược xuôi dòng.

(2) Thu được mẫu ADN bổ trợ bằng cách tiến hành phản ứng phiên mã ngược đối với ARN tổng số: 1µg ARN tổng số được phiên mã ngược theo quy trình của kit Master Mix SuperScript® VILO™ (Invitrogen).

(3) TLDA đối với ADN bổ trợ: các sản phẩm ADN bổ trợ nêu trên được trộn kỹ với hỗn hợp Master Mix dùng cho phản ứng PCR phổ dụng TaqMan®, và sau đó tiến hành các thử nghiệm trên thiết bị phản ứng PCR định lượng huỳnh quang ABI 7900 theo quy trình chuẩn TLDA.

(4) Phân tích và xử lý số liệu:

Trong pha thứ nhất của nghiên cứu, tổng cộng 192 gen ứng viên liên quan đến bệnh ung thư vú được thử nghiệm bằng kỹ thuật TLDA. Sự khác biệt về mức độ biểu hiện gen giữa 26 mẫu bệnh ung thư vú có sự tiên lượng tốt và 26 mẫu bệnh ung thư vú có sự tiên lượng xấu được phát hiện, và các gen có mức độ biểu hiện khác nhau được sàng lọc. Sự khác biệt về mức độ biểu hiện gen được thể hiện dưới dạng $2^{-\Delta Ct}$, trong đó $\Delta Ct = CT_{mẫu} - CT_{tham chiếu}$, và các gen tham chiếu được sử dụng làm thông số tham chiếu để chuẩn hóa và tính toán mức độ biểu hiện tương đối. Quá trình sàng lọc gen tham chiếu

chuẩn nội bao gồm các bước: sàng lọc gen tham chiếu nội ứng viên bằng cách sử dụng các chương trình phần mềm geNorm, BestKeeper, Normfinder, delta Ct trên cơ sở thuật toán ổn định và xem xét chức năng sinh học của các gen ít biến động và mối liên quan của chúng với khối u; tính toán sự tương quan giữa giá trị Ct trung bình của tất cả các tổ hợp gen tham chiếu chuẩn nội ứng viên và giá trị Ct trung bình của 192 gen, và tổ hợp có sự tương quan cao nhất là tổ hợp gen tham chiếu chuẩn nội, chứa các gen *ACTB*, *GAPDH*, *GUSB*, *NUP214* và *VCAN*. Tiêu chuẩn sàng lọc của gen ứng viên bao gồm (1) phân tích tổng quan- mức độ khác biệt giữa hai nhóm có sự tiên lượng tốt và sự tiên lượng xấu bằng 2 lần hoặc nhỏ hơn 0,5, và tỷ lệ các ca bệnh có giá trị Ct < 35 bằng 50%; (2) phân tích phân tầng – liên quan đến sự không di căn hạch bạch huyết, mức độ khác biệt giữa hai nhóm có sự tiên lượng tốt và sự tiên lượng xấu là gấp hơn 2 lần, và sự khác biệt về mặt thống kê là nhỏ hơn <0,05; (3) trong phép phân tích tổng quan, mức độ khác biệt giữa hai nhóm là không đáng kể nhưng đã được thông báo trong tài liệu chuyên ngành liên quan đến sự tiên lượng bệnh ung thư vú, và tỷ lệ các ca bệnh có giá trị Ct < 35 bằng 90%. Các gen đáp ứng tiêu chí sàng lọc trên đây bao gồm *BCL2*, *PGR*, *SCUBE2*, *ESR1*, *MKi67*, *CCNB1*, *MYBL2*, *GRB7*, *ERBB2*, *MMP11*, *CD68*, *BAG1*, *MAPT* và *MS4A1*. Chức năng của 14 gen nêu trên được thể hiện trong Bảng 1 sau đây.

Bảng 1: phân tích chức năng gen

Số thứ tự	Tên gen	Sự liên quan về chức năng
1.	<i>BCL2</i>	Liên quan đến estrogen
2.	<i>PGR</i>	Liên quan đến estrogen
3.	<i>SCUBE2</i>	Liên quan đến estrogen
4.	<i>ESR1</i>	Liên quan đến estrogen
5.	<i>MKi67</i>	Liên quan đến sự tăng sinh
6.	<i>MKi67</i>	Liên quan đến sự tăng sinh
7.	<i>MYBL2</i>	Liên quan đến sự tăng sinh
8.	<i>GRB7</i>	Liên quan đến Her-2
9.	<i>ERBB2</i>	Liên quan đến Her-2
10.	<i>MMP11</i>	Liên quan đến sự xâm lấn
11.	<i>CD68</i>	Họ biệt hóa 68
12.	<i>BAG1</i>	gen 1 kháng lại sự chết tế bào theo chương trình gắn

		kết với BCL2
13.	<i>MAPT</i>	protein tau liên quan đến vi ống
14.	<i>MS4A1</i>	Thành viên 1, siêu họ A của vùng xuyên màng 4

Ví dụ 3: kiểm tra các gen đánh dấu phân tử bằng phản ứng qRT-PCR với kích thước mẫu lớn

14 gen đánh dấu phân tử và 5 gen tham chiếu nội được sàng lọc bằng TLDA: *BCL2*, *PGR*, *SCUBE2*, *ESR1*, *MKI67*, *CCNB1*, *MYBL2*, *GRB7*, *ERBB2*, *MMP11*, *CD68*, *BAG1*, *MAPT*, *MS4A1*; *ACTB*, *GAPDH*, *GUSB*, *NUP214* và *VCAN*. Tổng cộng 289 mẫu FFPE đáp ứng các yêu cầu về mẫu nêu trên và có thông tin theo dõi trong lâm sàng đầy đủ được kiểm tra bằng phản ứng qRT-PCR trong một ống.

(1) Chiết ARN từ 289 mẫu FFPE: 4 lát cắt có độ dày 20µm hoặc 8 lát cắt có độ dày 10µm của mỗi mẫu được lấy; ARN được chiết theo quy trình của kit phân lập FFPE có độ tinh khiết cao (Roche); và ARN chiết được đạt sự kiểm soát chất lượng và số lượng bằng cách sử dụng NanoDrop-2000 được cho thử nghiệm phiên mã ngược xuôi dòng.

(2) Thu được mẫu ADN bổ trợ của 289 mẫu FFPE bằng cách tiến hành phản ứng phiên mã ngược đối với ARN tổng số: 1µg ARN tổng số được phiên mã ngược theo quy trình của kit Master Mix SuperScript® VILO™ (Invitrogen).

(3) Phản ứng qPCR đối với ADN bổ trợ từ 289 mẫu FFPE: các mẫu ADN bổ trợ trên đây được trộn kỹ với hỗn hợp Master Mix dùng cho phản ứng PCR phổ dụng TaqMan®, và sau đó tiến hành các thử nghiệm trên thiết bị TCR định lượng huỳnh quang ABI 7900 theo quy trình chuẩn của TLDA. Trình tự của các đoạn môi và đoạn dò dùng cho qPCR được thể hiện trong các bảng từ 2 đến 5.

Bảng 2: trình tự của các đoạn môi dùng cho qRT-PCR

Tên gen	Số	Đoạn môi ngược (5'-3')	Số	Đoạn môi xuôi (5'-3')
<i>BCL2</i>	SEQ ID	CATGTGTGTGGAGA	SEQ ID	GCCGGTTCAGGTACT
	NO:1	GCGTCAA	NO:2	CAGTCA
<i>PGR</i>	SEQ ID	CAGGTCTACCCGCC	SEQ ID	GGCTCTGGCTGGCTT
	NO:3	CTATCTC	NO:4	CTG
<i>SCUBE</i> 2	SEQ ID	AAGGATCACGGCTG	SEQ ID	AACCAGGCCTGCAC
	NO:5	TAGTCACA	NO:6	TCACA

<i>ESR1</i>	SEQ ID	TGATCAGGTCCACC	SEQ ID	GCGCCAGACGAGAC
	NO:7	TTCTAGAATG	NO:8	CAAT
<i>MKi67</i>	SEQ ID	GCACACTCCACCTG	SEQ ID	CCTGACTCTTGTTTT
	NO:9	TCCTGAA	NO:10	CCTGATGGT
<i>MKi67</i>	SEQ ID	CCAAGCCCAATGGA	SEQ ID	ATTACATCAGAGAA
	NO:11	AACATCT	NO:12	AGCCTGACACA
<i>MYBL2</i>	SEQ ID	AAGACAACAGCTTG	SEQ ID	GGGCCACTGCTGCCT
	NO:13	CTCAACCA	NO:14	TCT
<i>GRB7</i>	SEQ ID	TCCAGGACAAGGGC	SEQ ID	ATGGCGTCTGAGCG
	NO:15	ACACA	NO:16	AGAGA
<i>ERBB2</i>	SEQ ID	CCGAGGGCCGGTAT	SEQ ID	ATCCCACGTCCGTAG
	NO:17	ACATT	NO:18	AAAGGT
<i>MMP11</i>	SEQ ID	GATGCCCAGGGCCA	SEQ ID	TTCACCGTCGTACAC
	NO:19	CAT	NO:20	CCAGTAC
<i>CD68</i>	SEQ ID	AAGTTTCTCCTGCCC	SEQ ID	AAGAAGGATCAGGC
	NO:21	CAGTGA	NO:22	CGATGAT
<i>BAG1</i>	SEQ ID	GCTGGGCTCACCGT	SEQ ID	GAGGTAACATGAAG
	NO:23	GACT	NO:24	GTCGTGCTT
<i>MAPT</i>	SEQ ID	ACCAAGACCAAGAG	SEQ ID	CAGAGCCCCGGTTCCT
	NO:25	GGTGACA	NO:26	CAGA
<i>MS4A1</i>	SEQ ID	AACTCCCCATCTACC	SEQ ID	ATCACTGACAAAAT
	NO:27	CAATACTGTT	NO:28	GCCCAAGA

Bảng 3: các trình tự của đoạn dò dùng cho qRT-PCR

Tên gen	Số trình tự	Trình tự của đoạn dò
<i>BCL2</i>	SEQ ID NO:29	TGGACAACATCGCCCTG
<i>PGR</i>	SEQ ID NO:30	ACTACCTGAGGCCGGAT
<i>SCUBE2</i>	SEQ ID NO:31	TGCAAGGAGGCCC
<i>ESR1</i>	SEQ ID NO:32	CTGGCTAGAGATCCTGAT
<i>MKi67</i>	SEQ ID NO:33	TCATCAAGGAACAGCCTC
<i>MKi67</i>	SEQ ID NO:34	ATGTGCCCTGCAGAA

<i>MYBL2</i>	SEQ ID NO:35	CTTCTTGCAGGCCAAG
<i>GRB7</i>	SEQ ID NO:36	CTGGTTCCGTTAAGCC
<i>ERBB2</i>	SEQ ID NO:37	CGCCAGCTGTGTGAC
<i>MMP11</i>	SEQ ID NO:38	TGGTTCTTCCAAGGTGCT
<i>CD68</i>	SEQ ID NO:39	CCATCTTGCTGCCTCT
<i>BAG1</i>	SEQ ID NO:40	TCACCCACAGCAATGA
<i>MAPT</i>	SEQ ID NO:41	CCTGAAAGAATCTCC
<i>MS4A1</i>	SEQ ID NO:42	CAGCATACAATCTCTG

Bảng 4: các trình tự đoạn mồi của gen giữ nhà dùng cho qRT-PCR

Tên gen	Số	Đoạn mồi ngược (5'-3')	Số	Đoạn mồi xuôi (5'-3')
<i>ACTB</i>	SEQ ID NO:43	AGGCACCAGGGCGT GAT	SEQ ID NO:44	CTCGTCGCCCACATAG GAA
<i>GAPD</i>	SEQ ID NO:45	CAACGGATTTGGTC GTATTGG	SEQ ID NO:46	TGATGGCAACAATATC CACTTTACC
<i>H</i>				
<i>GUSB</i>	SEQ ID NO:47	TGAAGATGGTGATC GCTCACA	SEQ ID NO:48	GTTAGAGTTGCTCACA AAGGTCACA
<i>NUP214</i>	SEQ ID NO:49	ATCAGGCACAGGAG GGTTC A	SEQ ID NO:50	CCACCAAAACCCTGGA CAGA
<i>VCAN</i>	SEQ ID NO:51	ACTACAGGGCGGCA ACCA	SEQ ID NO:52	AACGTCCAAACAAGC CTTCTGA

Bảng 5: trình tự đoạn dò của gen giữ nhà dùng cho qRT-PCR

Tên gen	Số trình tự	Trình tự đoạn dò
<i>ACTB</i>	SEQ ID NO:53	TGGGCATGGGTCAGAA
<i>GAPDH</i>	SEQ ID NO:54	AGGGCTGCTTTTAACTC
<i>GUSB</i>	SEQ ID NO:55	ATCCTTGGACCCCTCC
<i>NUP214</i>	SEQ ID NO:56	CTTTGGGTCAAATAAC
<i>VCAN</i>	SEQ ID NO:57	CACACTGAATTTTGAGGCTG

Ví dụ 4: thiết lập mô hình dự đoán nguy cơ tái phát hoặc tử vong trong thời gian từ 3 đến 10 năm sau khi phẫu thuật ung thư vú

Thuật toán rừng ngẫu nhiên trong phương pháp học bằng máy được sử dụng để đánh giá và thử nghiệm giá trị về nguy cơ tái phát hoặc tử vong của các mẫu trong thời gian từ 3 đến 10 năm sau khi phẫu thuật, và thiết lập mô hình thử nghiệm gen để đánh giá việc chẩn đoán bệnh ung thư vú. Thuật toán rừng ngẫu nhiên gồm nhiều cây quyết định. Cây quyết định được xây dựng bằng phương pháp chọn ngẫu nhiên hai lần đối với thuộc tính và mẫu, vì thế nó còn được gọi là cây quyết định ngẫu nhiên. Trong thuật toán rừng ngẫu nhiên, không có sự tương quan giữa các cây quyết định. Khi số liệu thử nghiệm được nhập vào rừng ngẫu nhiên, nó được phân loại bởi mỗi cây quyết định. Cuối cùng, danh mục có sự phân loại nhiều nhất trong tất cả các cây quyết định là kết quả cuối cùng, nghĩa là kết quả “lựa chọn” của cây quyết định. Nói theo cách khác, thuật toán rừng ngẫu nhiên là công cụ phân loại bao gồm nhiều cây quyết định, và danh mục đầu ra của nó được xác định theo phương thức của danh mục đầu ra của cây riêng biệt. Theo sáng chế, tác giả sáng chế đã tối ưu hóa thuật toán rừng ngẫu nhiên truyền thống. Tác giả sáng chế chọn mẫu theo ngẫu nhiên 1000 lần và thiết lập 1000 mô hình. 39 mô hình tối ưu có cả độ chính xác và độ đặc hiệu cao từ 1000 mô hình được chọn làm mô hình con của mô hình cuối, và giá trị trung bình của 39 mô hình con được thiết lập làm kết quả dự đoán cuối.

Ví dụ 5: kiểm chứng thêm bằng các mẫu lâm sàng độc lập

19 mẫu FFPE từ các bệnh nhân ung thư vú dương tính với thụ thể ER hoặc PR giai đoạn trung gian hoặc giai đoạn sớm có số liệu theo dõi lâm sàng đã biết: giai đoạn PT là giai đoạn 1 và 2, các lần phẫu thuật được thực hiện trong khoảng thời gian từ năm 2004 đến năm 2008, tiến hành theo dõi từ năm 2011 đến năm 2015 và khoảng thời gian theo dõi là từ 3 đến 10 năm.

ARN tổng số của 19 mẫu FFPE nêu trên được chiết bằng cách sử dụng kit tách chiết ARN FFPET có độ tinh khiết cao (Roche). Sau khi việc kiểm soát chất lượng đủ điều kiện, ARN được cho phản ứng phiên mã ngược để thu được mẫu ADN bổ trợ. Các mẫu ADN bổ trợ được cho phản ứng qRT-PCR để phát hiện gen tham chiếu nội *ACTB*, *GAPDH*, *GUSB*, *NUP214*, *VCAN*, và genes *BCL2*, *PGR*, *SCUBE2*, *ESR1*, *MKI67*, *CCNB1*, *MYBL2*, *GRB7*, *ERBB2*, *MMP11*, *CD68*, *BAG1*, *MAPT* và *MS4A1*. Các giá trị Ct được đưa vào mô hình dự đoán nguy cơ tái phát hoặc tử vong trong thời gian từ 3 đến

10 năm sau khi phẫu thuật ung thư vú được xây dựng bằng thuật toán rừng ngẫu nhiên và thu được các giá trị về nguy cơ tái phát hoặc tử vong trong thời gian từ 3 đến 10 năm. Kết quả về sự tiên lượng tốt hoặc sự tiên lượng xấu được dự đoán trên cơ sở ngưỡng nguy cơ. Tỷ lệ phù hợp giữa kết quả được dự đoán và thông tin theo dõi đã biết là 73,6%. Các kết quả chi tiết được thể hiện trong Bảng 6.

Bảng 6: Kết quả đánh giá sự tiên lượng của 19 bệnh nhân ung thư vú

Mẫu số	Ngày phẫu thuật	Ngày theo dõi	Thời gian tái phát và di căn	Kết quả dự đoán
43200	2004-6-1	2015-11-6	Không tái phát trong quá trình theo dõi	Sự tiên lượng tốt
43346	2004-8-1	2010-5-1	Không tái phát trong quá trình theo dõi	Sự tiên lượng xấu
45110	2004-12-1	2011-1-1	06/2008: Di căn gan	Sự tiên lượng xấu
50170	2005-1-13	2014-11-1	Không tái phát trong quá trình theo dõi	Sự tiên lượng xấu
53552	2005-8-1	2010-2-13	9/2006: caxinom tuyến được biệt hóa ở mức thấp tái phát ở thành ngực trái; 10/2006: di căn phổi phải; 02/2009: di căn cả hai phổi	Sự tiên lượng xấu
53827	2005-8-18	2014-11-1	Không tái phát trong quá trình theo dõi	Sự tiên lượng tốt
54564	2005-9-28	2009-08-28	12/2006: caxinom tuyến được biệt hóa ở mức thấp tái phát ở nách phải; 12/2007: có sự di căn trung thất, phổi phải, hai bên cổ sau khi phẫu thuật ung thư vú ở bên phải	Sự tiên lượng xấu
54674	2005-10-10	2014-11-1	Không tái phát trong quá	Sự tiên

			trình theo dõi	lượng tốt
54775	2005-10-13	2007-7-27	6/6/2006: caxinom tuyến được biệt hóa ở mức thấp tái phát; 11/2006: nách phải, hai bên gian xương bả vai, màng phổi trái	Sự tiên lượng xấu
54938	2005-10-20	2014-11-1	Không tái phát trong quá trình theo dõi	Sự tiên lượng xấu
55335	2005-11-11	2008-11-19	12/2007: Di căn gan	Sự tiên lượng xấu
55475	2005-11-22	2012-4-12	10/12/2009: Di căn gan; 28/10/2010: Di căn gan và di căn hạch bạch huyết cổ tử cung phải; 11/11/2011: sự di căn xương	Sự tiên lượng tốt
55706	2005-12-8	2014-11-1	Không tái phát trong quá trình theo dõi	Sự tiên lượng tốt
55709	2005-12-12	2015-11-6	2/2009: caxinom ống tuyến vú xâm lấn tái phát liên quan đến cơ ngực; 2/2009: di căn xương; 6/2009: sự di căn hạch bạch huyết nách phải	Sự tiên lượng xấu
55792	2005-12-15	2009-4-1	01/2008: di căn đa xương	Sự tiên lượng xấu
60546	2006-2-16	2013-6-1	Không tái phát trong quá trình theo dõi	Sự tiên lượng tốt
70466	2007-1-29	2014-11-1	Không tái phát trong quá trình theo dõi	Sự tiên lượng tốt
76032	2007-11-16	2014-11-1	Không tái phát trong quá trình theo dõi	Sự tiên lượng tốt

86967	2008-12-25	2015-3-29	Không tái phát trong quá trình theo dõi	Sự tiên lượng tốt
-------	------------	-----------	---	-------------------

Bảng 7

Kết quả kiểm tra từ 19 mẫu độc lập						
Mẫu	Dự đoán bệnh lý			Hệ thống thử nghiệm để đánh giá sự tiên lượng bệnh ung thư vú		
Loại	Sự tiên lượng tốt	Sự tiên lượng xấu	Kết quả dự đoán	Sự tiên lượng tốt	Sự tiên lượng xấu	Kết quả dự đoán
Sự tiên lượng tốt	7	4	72,70%	8	3	72,70%
(theo dõi)			(độ đặc hiệu)			(độ đặc hiệu)
Sự tiên lượng xấu	1	7	87,50%	1	7	87,50%
(theo dõi)			(độ nhạy)			(độ nhạy)

Ví dụ 6

Hiện nay, việc quyết định điều trị lâm sàng và kế hoạch điều trị đối với bệnh ung thư vú phụ thuộc chủ yếu vào kết quả khám bệnh lý. Đồng thời, kết quả của việc khám bệnh lý là cơ sở khách quan quan trọng nhất để dự đoán sự tiên lượng. Hệ thống thử nghiệm sử dụng 289 mẫu FFPE từ các bệnh nhân ung thư vú mới được chẩn đoán có số liệu theo dõi trong lâm sàng đã biết được thu thập bởi Bệnh viện ung thư Đại học y Thiên Tân và Bệnh viện ung thư Henan được thử nghiệm bằng hệ thống đánh giá của sáng chế đối với 5 gen tham chiếu nội và 14 gen đánh dấu phân tử.

Các kết quả được thể hiện trong Bảng 8.

Bảng 8

Kết quả kiểm chứng từ 289 (kích thước mẫu lớn)		
Mẫu	Dự đoán bệnh lý	Hệ thống thử nghiệm đánh giá để tiên lượng bệnh ung thư vú

Loại	Sự tiên lượng tốt	Sự tiên lượng xấu	Kết quả dự đoán	Sự tiên lượng tốt	Sự tiên lượng xấu	Kết quả dự đoán
Sự tiên lượng tốt	128	50	71,90%	142	33	81,14%
(theo dõi)			(độ đặc hiệu)			(độ đặc hiệu)
Sự tiên lượng xấu	48	63	56,77%	52	62	54,39%
(theo dõi)			(độ nhạy)			(độ nhạy)

Kit theo sáng chế cho tỷ lệ chính xác 81,1% đối với các bệnh nhân ung thư vú mới được chẩn đoán nguy cơ tái phát hoặc tử vong thấp trong thời gian từ 3 đến 10 năm, trong khi tỷ lệ chính xác của thử nghiệm bệnh lý là 71,9%. Độ nhạy của tỷ lệ chính xác được cung cấp bởi kit này là 54,4% đối với các bệnh nhân ung thư vú mới được chẩn đoán nguy cơ tái phát hoặc tử vong cao trong thời gian từ 3 đến 10 năm, kết quả này gần bằng tỷ lệ chính xác của việc dự đoán bệnh lý tương ứng bằng 56,8%. So với thông tin theo dõi trong lâm sàng, kit này đạt được tỷ lệ phù hợp bằng 70%. Ngoại trừ độ tuổi, giai đoạn pT, và số lượng LN của các bệnh nhân, không cần phải dựa vào các thông tin bệnh lý lâm sàng khác.

Hệ thống thử nghiệm và kit này có tính năng thử nghiệm đánh giá tiên lượng bệnh ung thư vú tốt hơn so với các kết quả đánh giá bệnh lý lâm sàng. Nó có thể làm giảm ở mức độ nhất định đối với việc điều trị quá mức và điều trị không thích hợp do sự chẩn đoán bệnh lý không chính xác, đáp ứng các yêu cầu của việc điều trị được cá nhân hóa và chính xác của các bệnh nhân ung thư vú và cải thiện thêm các phương pháp kỹ thuật để dự đoán tiên lượng bệnh ung thư vú ở Trung Quốc.

Các gen đánh dấu phân tử và ứng dụng của chúng, kit thử nghiệm và phương pháp xây dựng mô hình thử nghiệm được đề xuất bởi sáng chế đã được mô tả chi tiết trên

đây. Các nguyên lý và phương án của sáng chế đã được mô tả cùng với các ví dụ cụ thể, và phần mô tả các phương án trên đây chỉ nhằm trợ giúp cho dễ hiểu phương pháp của sáng chế và ý tưởng cốt lõi của nó. Cần lưu ý rằng người có hiểu biết trung bình trong lĩnh vực này có thể tiến hành nhiều thay đổi và cải biến sáng chế mà vẫn nằm trong nguyên lý của sáng chế, và nằm trong phạm vi bảo hộ của sáng chế.

YÊU CẦU BẢO HỘ

1. Bộ đoạn mồi để khuếch đại tổ hợp gen, trong đó

bộ đoạn mồi này có trình tự được thể hiện trong các trình tự từ SEQ ID NO: 1 đến SEQ ID NO: 28; và

tổ hợp gen chứa các gen *BCL2*, *PGR*, *SCUBE2*, *ESR1*, *MKi67*, *CCNB1*, *MYBL2*, *GRB7*, *ERBB2*, *MMP11*, *CD68*, *BAG1*, *MAPT* và *MS4A1*.

2. Bộ đoạn dò để khuếch đại tổ hợp gen, trong đó

bộ đoạn dò này có trình tự được thể hiện trong các trình tự từ SEQ ID NO: 29 đến SEQ ID NO: 42; và

tổ hợp gen chứa các gen *BCL2*, *PGR*, *SCUBE2*, *ESR1*, *MKi67*, *CCNB1*, *MYBL2*, *GRB7*, *ERBB2*, *MMP11*, *CD68*, *BAG1*, *MAPT* và *MS4A1*.

3. Kit thử nghiệm để dự đoán nguy cơ tái phát và/hoặc tử vong trong thời gian từ 3 đến 10 năm sau khi phẫu thuật ung thư vú, trong đó kit này chứa bộ đoạn mồi theo điểm 1 và bộ đoạn dò theo điểm 2, và các chất phản ứng thường được sử dụng trong kit.

DANH MỤC TRÌNH TỰ

<110> CAPITALBIO CORPORATION; BEIJING CAPITALBIO MEDLAB CO., LTD.

<120> BỘ ĐOẠN MỖI VÀ BỘ ĐOẠN DÒ ĐỀ KHUẾCH ĐẠI TỔ HỢP GEN VÀ
GEN THAM CHIẾU NỘI, KIT THỬ NGHIỆM CHỨA CHỨNG

<130> OP170656

<160> 57

<170> PatentIn version 3.3

<210> 1

<211> 21

<212> ADN

<213> Trình tự nhân tạo

<400> 1

catgtgtgtg gagagcgtca a 21

<210> 2

<211> 21

<212> ADN

<213> Trình tự nhân tạo

<400> 2

gccggttcag gtactcagtc a 21

<210> 3

<211> 21

<212> ADN

<213> Trình tự nhân tạo

<400> 3

caggtctacc cgccctatct c 21

<210> 4

<211> 18

<212> ADN

<213> Trình tự nhân tạo

<400> 4

ggctctggct ggcttctg 18

<210> 5
 <211> 22
 <212> ADN
 <213> Trình tự nhân tạo
 <400> 5
 aaggatcacg gctgtagtca ca 22
 <210> 6
 <211> 19
 <212> ADN
 <213> Trình tự nhân tạo
 <400> 6
 aaccaggcct gcactcaca 19
 <210> 7
 <211> 24
 <212> ADN
 <213> Trình tự nhân tạo
 <400> 7
 tgatcaggtc caccttctag aatg 24
 <210> 8
 <211> 18
 <212> ADN
 <213> Trình tự nhân tạo
 <400> 8
 gcgccagacg agaccaat 18
 <210> 9
 <211> 21
 <212> ADN
 <213> Trình tự nhân tạo
 <400> 9
 gcacactcca cctgtcctga a 21
 <210> 10

<211> 24
 <212> ADN
 <213> Trình tự nhân tạo
 <400> 10
 cctgactcctt gttttcctga tgggt 24
 <210> 11
 <211> 21
 <212> ADN
 <213> Trình tự nhân tạo
 <400> 11
 ccaagcccaa tggaaacatc t 21
 <210> 12
 <211> 25
 <212> ADN
 <213> Trình tự nhân tạo
 <400> 12
 attacatcag agaaagcctg acaca 25
 <210> 13
 <211> 22
 <212> ADN
 <213> Trình tự nhân tạo
 <400> 13
 aagacaacag cttgctcaac ca 22
 <210> 14
 <211> 18
 <212> ADN
 <213> Trình tự nhân tạo
 <400> 14
 gggccactgc tgccttct 18
 <210> 15
 <211> 19

<212> ADN
 <213> Trình tự nhân tạo
 <400> 15
 tccaggacaa gggcacaca 19
 <210> 16
 <211> 19
 <212> ADN
 <213> Trình tự nhân tạo
 <400> 16
 atggcgtctg agcgagaga 19
 <210> 17
 <211> 19
 <212> ADN
 <213> Trình tự nhân tạo
 <400> 17
 ccgagggccg gtatacatt 19
 <210> 18
 <211> 21
 <212> ADN
 <213> Trình tự nhân tạo
 <400> 18
 atcccacgtc cgtagaaagg t 21
 <210> 19
 <211> 17
 <212> ADN
 <213> Trình tự nhân tạo
 <400> 19
 gatgcccagg gccacat 17
 <210> 20
 <211> 22
 <212> ADN

<213> Trình tự nhân tạo
 <400> 20
 ttcaccgtcg tacacccagt ac 22
 <210> 21
 <211> 21
 <212> ADN
 <213> Trình tự nhân tạo
 <400> 21
 aagtttctcc tgccccagtg a 21
 <210> 22
 <211> 21
 <212> ADN
 <213> Trình tự nhân tạo
 <400> 22
 aagaaggatc aggccgatga t 21
 <210> 23
 <211> 18
 <212> ADN
 <213> Trình tự nhân tạo
 <400> 23
 gctgggctca ccgtgact 18
 <210> 24
 <211> 23
 <212> ADN
 <213> Trình tự nhân tạo
 <400> 24
 gaggtaacat gaaggtcgtg ctt 23
 <210> 25
 <211> 21
 <212> ADN
 <213> Trình tự nhân tạo

<400>	25	
accaagacca agaggggtgac a		21
<210>	26	
<211>	19	
<212>	ADN	
<213>	Trình tự nhân tạo	
<400>	26	
cagagcccgg ttcctcaga		19
<210>	27	
<211>	25	
<212>	ADN	
<213>	Trình tự nhân tạo	
<400>	27	
aactcccat ctaccaata ctgtt		25
<210>	28	
<211>	22	
<212>	ADN	
<213>	Trình tự nhân tạo	
<400>	28	
atcactgaca aaatgcccaa ga		22
<210>	29	
<211>	17	
<212>	ADN	
<213>	Trình tự nhân tạo	
<400>	29	
tggaacaacat cgccctg		17
<210>	30	
<211>	17	
<212>	ADN	
<213>	Trình tự nhân tạo	
<400>	30	

actacctgag gccggat	17
<210> 31	
<211> 13	
<212> ADN	
<213> Trình tự nhân tạo	
<400> 31	
tgcaaggagg ccc	13
<210> 32	
<211> 18	
<212> ADN	
<213> Trình tự nhân tạo	
<400> 32	
ctggctagag atcctgat	18
<210> 33	
<211> 18	
<212> ADN	
<213> Trình tự nhân tạo	
<400> 33	
tcatcaagga acagcctc	18
<210> 34	
<211> 16	
<212> ADN	
<213> Trình tự nhân tạo	
<400> 34	
atgtgcccct gcagaa	16
<210> 35	
<211> 16	
<212> ADN	
<213> Trình tự nhân tạo	
<400> 35	
cttcttgag gccaa	16

<210>	36	
<211>	16	
<212>	ADN	
<213>	Trình tự nhân tạo	
<400>	36	
	ctggttccgt taagcc	16
<210>	37	
<211>	15	
<212>	ADN	
<213>	Trình tự nhân tạo	
<400>	37	
	cgccagctgt gtagc	15
<210>	38	
<211>	18	
<212>	ADN	
<213>	Trình tự nhân tạo	
<400>	38	
	tggttcttcc aaggtgct	18
<210>	39	
<211>	16	
<212>	ADN	
<213>	Trình tự nhân tạo	
<400>	39	
	ccatcttgct gcctct	16
<210>	40	
<211>	16	
<212>	ADN	
<213>	Trình tự nhân tạo	
<400>	40	
	tcacccacag caatga	16
<210>	41	

<211>	15	
<212>	ADN	
<213>	Trình tự nhân tạo	
<400>	41	
cctgaaagaa tctcc		15
<210>	42	
<211>	16	
<212>	ADN	
<213>	Trình tự nhân tạo	
<400>	42	
cagcatacaa tctctg		16
<210>	43	
<211>	17	
<212>	ADN	
<213>	Trình tự nhân tạo	
<400>	43	
aggcaccagg gcgtgat		17
<210>	44	
<211>	19	
<212>	ADN	
<213>	Trình tự nhân tạo	
<400>	44	
ctcgtcgccc acataggaa		19
<210>	45	
<211>	21	
<212>	ADN	
<213>	Trình tự nhân tạo	
<400>	45	
caacggatatt ggtcgtattg g		21
<210>	46	
<211>	25	

<212> ADN
 <213> Trình tự nhân tạo
 <400> 46
 tgatggcaac aatatccact ttacc 25
 <210> 47
 <211> 21
 <212> ADN
 <213> Trình tự nhân tạo
 <400> 47
 tgaagatggg gatcgctcac a 21
 <210> 48
 <211> 25
 <212> ADN
 <213> Trình tự nhân tạo
 <400> 48
 gttagagttg ctcacaaagg tcaca 25
 <210> 49
 <211> 20
 <212> ADN
 <213> Trình tự nhân tạo
 <400> 49
 atcaggcaca ggagggttca 20
 <210> 50
 <211> 20
 <212> ADN
 <213> Trình tự nhân tạo
 <400> 50
 ccaccaaac cctggacaga 20
 <210> 51
 <211> 18
 <212> ADN

<213> Trình tự nhân tạo

<400> 51

actacagggc ggcaacca

18

<210> 52

<211> 22

<212> ADN

<213> Trình tự nhân tạo

<400> 52

aacgtccaaa caagccttct ga

22

<210> 53

<211> 16

<212> ADN

<213> Trình tự nhân tạo

<400> 53

tgggcatggg tcagaa

16

<210> 54

<211> 17

<212> ADN

<213> Trình tự nhân tạo

<400> 54

agggctgctt ttaactc

17

<210> 55

<211> 16

<212> ADN

<213> Trình tự nhân tạo

<400> 55

atccttggac ccctcc

16

<210> 56

<211> 16

<212> ADN

<213> Trình tự nhân tạo

<400> 56

ctttgggtca aataac

16

<210> 57

<211> 20

<212> ADN

<213> Trình tự nhân tạo

<400> 57

cacactgaat tttgaggctg

20

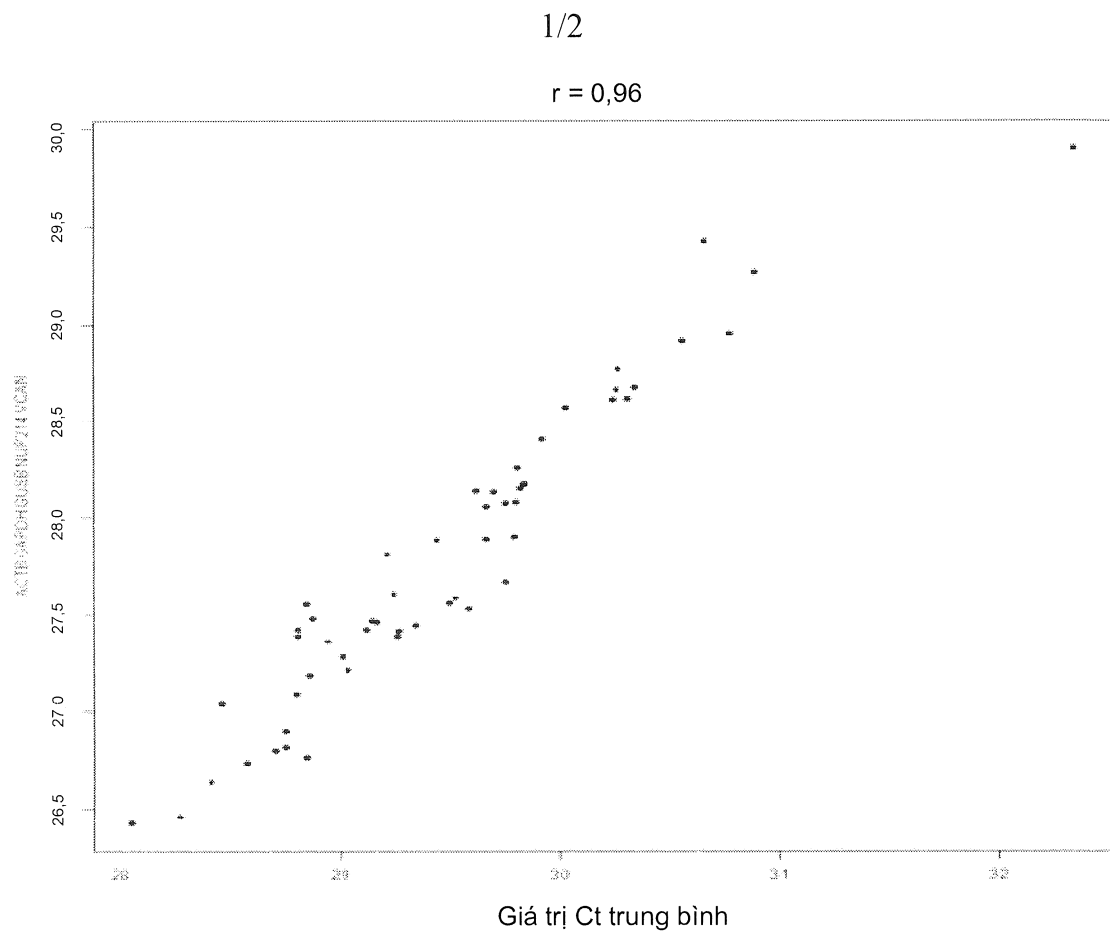


Fig.1

2/2

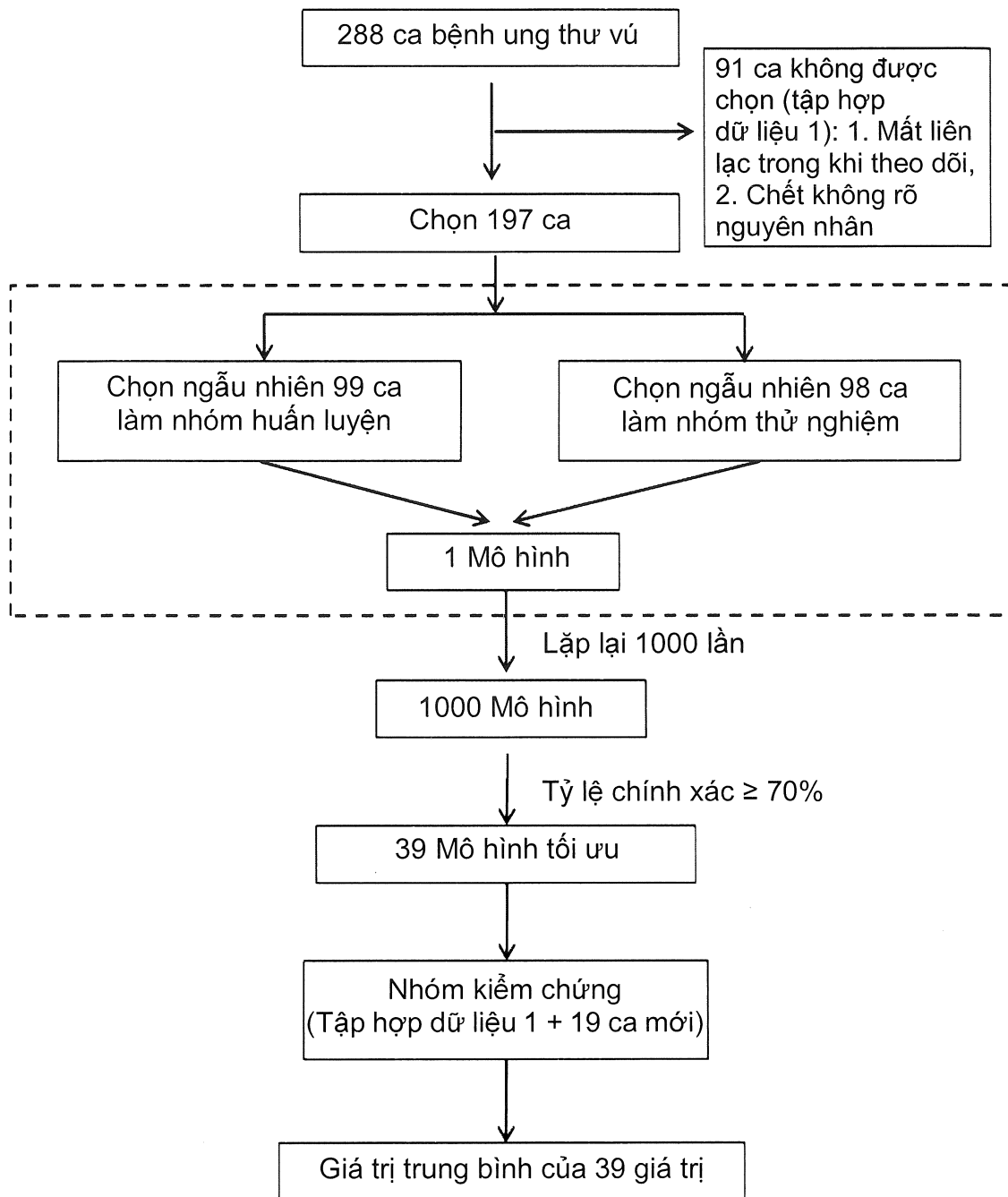


Fig.2