



(12)发明专利

(10)授权公告号 CN 103492566 B

(45)授权公告日 2017.06.13

(21)申请号 201280020542.4

(22)申请日 2012.04.25

(65)同一申请的已公布的文献号  
申请公布号 CN 103492566 A

(43)申请公布日 2014.01.01

(30)优先权数据  
097432/2011 2011.04.25 JP

(85)PCT国际申请进入国家阶段日  
2013.10.25

(86)PCT国际申请的申请数据  
PCT/JP2012/061100 2012.04.25

(87)PCT国际申请的公布数据  
W02012/147800 JA 2012.11.01

(73)专利权人 东丽株式会社  
地址 日本东京都  
专利权人 国立大学法人京都大学

(72)发明人 妙本阳 小园聪子 佐藤史显  
户井雅和 上野贵之 王智鹏  
辻本豪三 清水一治

(74)专利代理机构 北京市中咨律师事务所  
11247

代理人 曾祯 段承恩

(51)Int.Cl.  
C12N 15/09(2006.01)  
C12Q 1/68(2006.01)  
A61K 39/395(2006.01)

(56)对比文件  
WO 2006/118308 A1,2006.11.09,  
WO 2007/026896 A1,2007.03.08,  
JP 2008/535482 A,2008.09.04,  
WO 2009/148593 A1,2009.12.10,  
US 2011/0124521 A1,2011.05.26,  
Ve<sup>1</sup>gran等.Gene expression profile and  
response to trastuzumab-docetaxel-based.  
《British Journal of Cancer》.2009,第101卷  
图1,表4-6.

审查员 张锦广

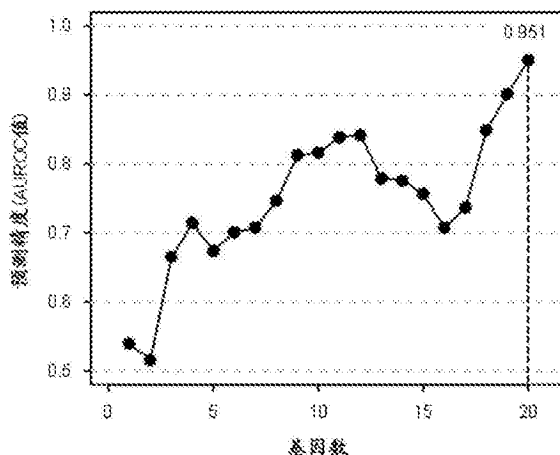
权利要求书2页 说明书30页  
序列表5页 附图3页

(54)发明名称

乳癌患者对曲妥珠单抗的治疗敏感性预测  
用组合物和方法

(57)摘要

本发明涉及用于乳癌患者对曲妥珠单抗的  
治疗敏感性预测的组合物和方法,具体地,涉及  
使用包含序列表所记载的序列号1~9、序列号11  
~19、和序列号21~23所示的碱基序列、或该碱  
基序列中u为t的碱基序列的多核苷酸、该多核苷  
酸的突变体、该多核苷酸的衍生物、或该多核苷  
酸的包含16个以上连续的碱基的片段、或者包含  
与它们互补的碱基序列的多核苷酸,以Her2蛋白  
质的表达量的增减为指标而用于预测对曲妥珠  
单抗的治疗敏感性的组合物,试剂盒,DNA芯片,  
以及方法。



1. 用于测定miR-26a的表达的选自下述(a)~(b)所示的多核苷酸或该多核苷酸的片段中的1种以上的多核苷酸,与用于测定let-7a、let-7b、miR-1234、miR-513a-5p、miR-494、let-7g、miR-940、miR-1470、miR-200c、let-7e、miR-1228、let-7c、miR-1229、miR-205、miR-145、miR-181a、miR-191、miR-92a、let-7d和miR-23a的表达的选自下述(a')~(b')所示的多核苷酸或该多核苷酸的片段中的1种以上的多核苷酸的用途,用于制造乳癌患者对曲妥珠单抗的治疗敏感性预测用组合物,

(a) 由序列号4所示的碱基序列、或该碱基序列中u为t的碱基序列组成的多核苷酸或该多核苷酸的由16个以上连续的碱基组成的片段;

(b) 由与序列号4所示的碱基序列、或该碱基序列中u为t的碱基序列互补的碱基序列组成的多核苷酸或该多核苷酸的由16个以上连续的碱基组成的片段,

(a') 由序列号1~3、5~6、7~9、序列号11~19、和序列号21~23所示的碱基序列、或该碱基序列中u为t的碱基序列组成的多核苷酸或该多核苷酸的由16个以上连续的碱基组成的片段;

(b') 由与序列号1~3、5~6、7~9、序列号11~19、和序列号21~23所示的碱基序列、或该碱基序列中u为t的碱基序列互补的碱基序列组成的多核苷酸或该多核苷酸的由16个以上连续的碱基组成的片段。

2. 根据权利要求1所述的用途,所述组合物进一步包含用于测定miR-125a-5p和/或miR-125b的表达的选自下述(c)~(d)所示的多核苷酸和/或该多核苷酸的片段中的1种或2种多核苷酸:

(c) 由序列号10和序列号20所示的碱基序列、或该碱基序列中u为t的碱基序列组成的多核苷酸或该多核苷酸的由16个以上连续的碱基组成的片段;

(d) 由与序列号10和序列号20所示的碱基序列、或该碱基序列中u为t的碱基序列互补的碱基序列组成的多核苷酸或该多核苷酸的由16个以上连续的碱基组成的片段。

3. 用于测定miR-26a的表达的选自权利要求1所述的(a)~(b)所示的多核苷酸或该多核苷酸的片段中的1种以上的多核苷酸,与用于测定let-7a、let-7b、miR-1234、miR-513a-5p、miR-494、let-7g、miR-940、miR-1470、miR-200c、let-7e、miR-1228、let-7c、miR-1229、miR-205、miR-145、miR-181a、miR-191、miR-92a、let-7d和miR-23a的表达的选自权利要求1所述的(a')~(b')所示的多核苷酸或该多核苷酸的片段中的1种以上的多核苷酸的用途,用于制造乳癌患者对曲妥珠单抗的治疗敏感性预测用试剂盒。

4. 根据权利要求3所述的用途,所述试剂盒进一步包含用于测定miR-125a-5p和/或miR-125b的表达的选自权利要求2所述的(c)~(d)所示的多核苷酸和/或该多核苷酸的片段中的1种或2种多核苷酸。

5. 根据权利要求3或4所述的用途,所述多核苷酸是,由序列号1~23的任一项所示的碱基序列、或该碱基序列中u为t的碱基序列组成的多核苷酸、由该多核苷酸的互补序列组成的多核苷酸、或这些多核苷酸的由16个以上连续的碱基组成的片段。

6. 根据权利要求3~5的任一项所述的用途,所述多核苷酸单独或任意组合地被包装于不同的容器中。

7. 用于测定miR-26a的表达的选自权利要求1所述的(a)~(b)所示的多核苷酸和/或该多核苷酸的片段中的1种以上的多核苷酸,与用于测定let-7a、let-7b、miR-1234、miR-

513a-5p、miR-494、let-7g、miR-940、miR-1470、miR-200c、let-7e、miR-1228、let-7c、miR-1229、miR-205、miR-145、miR-181a、miR-191、miR-92a、let-7d和miR-23a的表达的选自权利要求1所述的(a')~(b')所示的多核苷酸或该多核苷酸的片段中的1种以上的多核苷酸的用途,用于制造乳癌患者对曲妥珠单抗的治疗敏感性预测用DNA芯片。

8. 根据权利要求7所述的用途,所述DNA芯片进一步包含用于测定miR-125a-5p和/或miR-125b的表达的权利要求2所述的(c)~(d)所示的多核苷酸和/或该多核苷酸的片段多核苷酸的1种或2种多核苷酸。

## 乳腺癌患者对曲妥珠单抗的治疗敏感性预测用组合物和方法

### 技术领域

[0001] 本发明涉及乳腺癌患者对曲妥珠单抗 (trastuzumab) 的治疗敏感性的预测 (或者, 判定、评价、检测或诊断) 有用的组合物、利用了该组合物的乳腺癌患者对曲妥珠单抗的治疗敏感性的预测 (或者、判定、评价、检测或诊断) 方法、和利用了该组合物的乳腺癌患者对曲妥珠单抗的治疗敏感性的预测 (或者、判定、评价、检测或诊断) 用试剂盒。

### 背景技术

[0002] 乳腺癌是指乳房组织内的细胞恶性化而无秩序地增殖的疾患, 女性的罹患率在日本为25人~30人中1人, 欧美为8~10人中1人。此外, 已知虽然罹患率低, 但男性也会罹患乳腺癌。根据最近的研究已知, 乳腺癌包含具有不同的生物学的特性的多样集团, 对治疗的反应性和/或预后在各个集团的患者中是不同的。即, 根据利用DNA芯片分析的全面的表达基因分析显示, 乳腺癌大致可以分为5种分子亚型。但是, 在日常临床中, 多通过检测雌激素受体、孕酮受体、Her2蛋白质的表达而分类为4种亚型来研究治疗方针。乳腺癌的治疗原则上为外科疗法, 根据进行度和/或转移、全身状态、乳腺癌的分类而合并使用药物疗法、放射线疗法。特别是在药物疗法的实施中, 依据上述的乳腺癌亚型评价应施与对象患者的药物、选择适当的治疗方针是重要的 (非专利文献1)。

[0003] 在这些亚型中, 占乳腺癌的约25%的Her2阳性乳腺癌的恶性度、转移率高、预后不良, Her2阳性乳腺癌的治疗成绩的提高今后也会成为极其重要的课题。

[0004] 曲妥珠单抗 (商品名“赫赛汀” (注册商标、中外制药)) 是厚生劳动省认可的抗体药物, 通过与存在于Her2阳性乳腺癌的细胞表面的Her2蛋白质结合而发挥抗肿瘤效果。曲妥珠单抗是对Her2阳性乳腺癌使用的第一选择药。但是另一方面, 还已知Her2阳性乳腺癌患者中存在没有曲妥珠单抗的效果的曲妥珠单抗非敏感性乳腺癌患者、由于曲妥珠单抗的给药而产生心力衰竭、呼吸困难和/或变态反应这样的严重副作用的患者。现在的临床诊断中, 用于区别乳腺癌是否为Her2阳性的方法是利用免疫组织化学方法的Her2蛋白质的过表达的检测、和/或与Her2蛋白质对应的基因组上的基因扩增的检测, 但该方法不能判别上述的虽然为Her2阳性但显示曲妥珠单抗非敏感性的乳腺癌患者、和并发副作用的乳腺癌患者。

[0005] 具体的, 已知通过免疫组织化学方法确认了Her2蛋白质的过表达、接受了利用曲妥珠单抗单药的治疗的患者中对曲妥珠单抗显示敏感性的患者的比例为35%以下 (非专利文献2), 在通过利用免疫组织化学方法的Her2蛋白质的过表达的检测或与Her2蛋白质对应的基因组上的基因扩增的检测来区别Her2阳性的检查方法 (非专利文献3的检查方法) 中, 接受了对曲妥珠单抗组合了其他抗癌剂的治疗的患者中对曲妥珠单抗显示敏感性的患者的比例为65.2%以下 (非专利文献3)。即, 现在在临床现场利用的非专利文献3的检查方法的曲妥珠单抗敏感性预测的精度止于65.2%。

[0006] 另一方面, 用于乳腺癌治疗的药物疗法近年来取得了特别显著的发展, 渐渐可以根据癌的性质选择各种治疗药。在这样的状况下, 如果能够以现在用于区别Her2阳性的方法以上的高精度预测乳腺癌患者对曲妥珠单抗的治疗敏感性, 则可以使Her2阳性乳腺癌患者的治

疗法的选择更容易,实现药物治疗的效果的最大化和副作用的最小化。

[0007] 到目前为止关于Her2阳性乳癌患者对曲妥珠单抗的治疗敏感性的报告中,存在关于非专利文献4所记载的PTEN蛋白质的活性化、非专利文献5所记载的细胞周期蛋白E(Cyclin E)的基因扩增和/或过表达、非专利文献6所记载的利用miR-125a和/或miR-125b的Her2蛋白质表达的控制的记载。

[0008] 在非专利文献4中,通过免疫组织学方法评价Her2阳性乳癌患者的PTEN的表达量,显示抑制了PTEN的表达的细胞不易受到由曲妥珠单抗引起的增殖抑制,PTEN的表达量与Her2阳性乳癌患者的由曲妥珠单抗产生的病情进行的抑制相关。

[0009] 另外在非专利文献5中,通过免疫组织化学方法来评价Her2阳性乳癌患者对曲妥珠单抗的治疗敏感性患者中的细胞周期蛋白E蛋白质的表达量,显示在接受了通过曲妥珠单抗和其他抗癌剂的治疗的患者组中,细胞周期蛋白E蛋白质的表达量多的组中病情不进行的患者的比例高。

[0010] 另外,非专利文献6显示miR-125a和miR-125b的表达量增加使作为曲妥珠单抗的目标蛋白质的Her2蛋白质的表达量降低。

[0011] 另外,已知let-7a(专利文献1)、let-7b(专利文献1)、和miR-145(专利文献2)在乳癌患者中表达量降低,并且miR-200c(专利文献3)在乳癌患者中表达量上升。

[0012] 现有技术文献

[0013] 专利文献

[0014] 专利文献1:US2008/0076674A1

[0015] 专利文献2:日本特表2010-510964

[0016] 专利文献3:日本特表2010-504350

[0017] 非专利文献

[0018] 非专利文献1:日本乳癌学会编乳癌取扱い規約(乳癌处理规章)第16版2008年、日本乳癌学会编乳癌診療ガイドライン(乳癌诊疗指南)1.药物治疗2010年版2010年

[0019] 非专利文献2:C.L.Vogel等、2002年、Journal of Clinical Oncology、第20卷、p.719-726

[0020] 非专利文献3:A.U.Buzdar等、2005年、Journal of Clinical Oncology、第23卷、p3676-3685

[0021] 非专利文献4:Yoichi,N.等、2004年、Cancer Cell、第6卷、p.117-127

[0022] 非专利文献5:Maurizio,S.等、2011年、Proc Natl Acad Sci USA.,Early Edition、pnas.1014835108

[0023] 非专利文献6:Scott,GK.等、2007年、J Biol Chem.、第282卷、p.1479-1486

## 发明内容

[0024] 发明要解决的课题

[0025] 这样,现有技术中已知多个与对曲妥珠单抗的治疗敏感性显示结果上相关的表达量的基因、蛋白质,但任一种均未充分地发现作为预测Her2阳性乳癌患者对曲妥珠单抗的治疗敏感性的标志物的价值。

[0026] 非专利文献4的PTEN不可能根据PTEN的表达量在施与前预测每个乳癌患者对曲妥

珠单抗的治疗敏感性。而且,非专利文献5的细胞周期蛋白E不能根据细胞周期蛋白E蛋白质的表达量而预测每个乳癌患者对曲妥珠单抗的治疗敏感性。另外,非专利文献6的miR-125a和miR-125b,miR-125a和miR-125b的表达量增加与Her2阳性乳癌患者对曲妥珠单抗的治疗敏感性的关系不明,使用任一种均不能预测每个乳癌患者对曲妥珠单抗的治疗敏感性。因此,不能使用这些表达量为指标进行一般临床上的利用,迫切要求预测精度更高的乳癌患者对曲妥珠单抗的治疗敏感性的预测标志物。

[0027] 本发明的目的是提供在乳癌患者对曲妥珠单抗的治疗敏感性的预测(或者,判定、评价、检测或诊断)中有用的组合物、利用了该组合物的乳癌患者对曲妥珠单抗的治疗敏感性的预测(或者,判定、评价、检测或诊断)方法、和利用了该组合物的乳癌患者对曲妥珠单抗的治疗敏感性的预测(或者,判定、评价、检测或诊断)用试剂盒。

[0028] 用于解决课题的方法

[0029] 作为乳癌患者对曲妥珠单抗的敏感性的预测用的基因标志物探索方法,可列举:将通过对乳癌患者的利用曲妥珠单抗的治疗或利用曲妥珠单抗与其他抗癌剂并用的治疗而在检查时、治疗前、治疗中、或治疗后采取的来源于乳癌患者的组织、体液、或分泌物所含的基因、蛋白质、或代谢产物等的量,在有对曲妥珠单抗的治疗敏感性的患者和无对曲妥珠单抗的治疗敏感性的患者之间使用某些手段进行比较的方法。

[0030] 使用了DNA芯片的基因表达量分析近年来作为这样的标志物探索的方法特别被通用。DNA芯片上固定有利用了与数百至数万方种基因对应的碱基序列的探针。通过将被测样品添加于DNA芯片而使样品中的基因与探针结合,通过某些手段测定该结合量,从而得知被测样品中的基因量。与固定化于DNA芯片上的探针对应的基因的选择是自由的,另外使用检查时、治疗前、治疗中、或治疗后采取的来源于乳癌患者的组织、FFPE标本、体液、或分泌物等样品,通过比较该样品中的基因表达量,可以推定能够成为可以在乳癌的诊断中利用的标志物的基因群。

[0031] 为了解决上述课题,本发明者们用DNA芯片分析由从治疗前的乳癌患者采取的乳癌病变部的针活检组织得到的基因表达量,发现了成为能够预测乳癌患者对曲妥珠单抗的治疗敏感性的标志物的基因,进而发现它们的基因表达量在对曲妥珠单抗的治疗敏感性高的乳癌患者中,从乳癌患者得到的乳癌病变部的基因表达量减少、减低或增加、增大,从而完成了本发明。

[0032] 1.发明的概要

[0033] 本发明具有以下特征。

[0034] 本发明在第1方式中,提供乳癌患者对曲妥珠单抗的治疗敏感性预测用组合物,包含选自下述(a)~(j)所示的多核苷酸、该多核苷酸的突变体、或该多核苷酸的片段中的2种以上的多核苷酸:

[0035] (a)由序列号1~9、序列号11~19、和序列号21~23所示的碱基序列、或该碱基序列中u为t的碱基序列组成的多核苷酸、该多核苷酸的突变体、该多核苷酸的衍生物、或该多核苷酸的包含16个以上连续的碱基的片段;

[0036] (b)包含序列号1~9、序列号11~19、和序列号21~23所示的碱基序列、或该碱基序列中u为t的碱基序列的多核苷酸;

[0037] (c)由与序列号1~9、序列号11~19、和序列号21~23所示的碱基序列、或该碱基

序列中u为t的碱基序列互补的碱基序列组成的多核苷酸、该多核苷酸的突变体、该多核苷酸的衍生物、或该多核苷酸的包含16个以上连续的碱基的片段；

[0038] (d) 包含与序列号1~9、序列号11~19、和序列号21~23所示的碱基序列、或该碱基序列中u为t的碱基序列互补的碱基序列的多核苷酸；

[0039] (e) 与所述(a)~(d)的任一多核苷酸在严格条件下杂交的多核苷酸、或该多核苷酸的包含16个以上连续的碱基的片段；

[0040] (f) 由序列号10和序列号20所示的碱基序列、或该碱基序列中u为t的碱基序列组成的多核苷酸、该多核苷酸的突变体、该多核苷酸的衍生物、或包含16个以上连续的碱基的该多核苷酸的片段；

[0041] (g) 包含序列号10和序列号20所示的碱基序列、或该碱基序列中u为t的碱基序列的多核苷酸；

[0042] (h) 由与序列号10和序列号20所示的碱基序列、或该碱基序列中u为t的碱基序列互补的碱基序列组成的多核苷酸、该多核苷酸的突变体、该多核苷酸的衍生物、或该多核苷酸的包含16个以上连续的碱基的片段；

[0043] (i) 包含与序列号10和序列号20所示的碱基序列、或该碱基序列中u为t的碱基序列互补的碱基序列的多核苷酸；

[0044] (j) 与所述(f)~(i)的任一多核苷酸在严格条件下杂交的多核苷酸、或该多核苷酸的包含16个以上连续的碱基的片段。

[0045] 本发明在第2方式中,提供乳癌患者对曲妥珠单抗的治疗敏感性预测用试剂盒,包含上述的(a)~(e)所示的多核苷酸、该多核苷酸的突变体、该多核苷酸的衍生物、和/或该多核苷酸的片段的2种以上。

[0046] 在该实施方式中,上述试剂盒进一步包含上述的(f)~(j)所示的多核苷酸、该多核苷酸的突变体、该多核苷酸的衍生物、和/或该多核苷酸的片段的1种或2种。

[0047] 另外,提供上述的试剂盒,其中,上述多核苷酸是,由序列号1~23的任一项所示的碱基序列、或该碱基序列中u为t的碱基序列组成的多核苷酸、由该多核苷酸的互补序列组成的多核苷酸、与这些多核苷酸在严格条件下杂交的多核苷酸、或这些多核苷酸的包含16个以上连续的碱基的2以上的片段。

[0048] 进而在其他实施方式中,提供上述的试剂盒,其中,上述多核苷酸单独或任意组合地被包装于不同的容器中。

[0049] 本发明在第3方式中,提供乳癌患者对曲妥珠单抗的治疗敏感性预测用DNA芯片,包含上述的(a)~(e)所示的多核苷酸、该多核苷酸的突变体、该多核苷酸的衍生物、和/或该多核苷酸的片段的2种以上。

[0050] 在该实施方式中,上述DNA芯片进一步包含上述的(f)~(j)所示的多核苷酸、该多核苷酸的突变体、该多核苷酸的衍生物、和/或该多核苷酸的片段的1种或2种。

[0051] 本发明在第4方式中,提供用于乳癌患者对曲妥珠单抗的治疗敏感性预测的方法,包括:使用上述的组合物的多核苷酸,测定来源于乳癌患者的样品中的目标核酸的表达量的2种以上,体外预测、判定或评价乳癌患者对曲妥珠单抗的治疗敏感性。

[0052] 在该实施方式中,上述的方法使用本发明的第2方式所述的试剂盒。

[0053] 在其他实施方式中,上述的方法使用本发明的第3方式所述的DNA芯片。

[0054] 本发明进一步在第5方式中,提供用于乳癌患者对曲妥珠单抗的治疗敏感性预测的方法,包括以下工序:第1工序,使用上述的任一项所述的组合物、上述的任一项所述的试剂盒、上述的任一项所述的DNA芯片、或它们的组合,体外测定已知对曲妥珠单抗具有治疗敏感性的乳癌患者的多个样品中的目标核酸的表达量;第2工序,测定所述第1工序所得的该目标核酸的表达量,制成以由该目标核酸的表达量计算出的基因表达量为教师的判别式(支持向量机);第3工序,与第1工序同样地体外测定乳癌患者的手术时或活检检查时采取的样品中的该目标核酸的表达量;第4工序,在所述第2工序所得的判别式中代入由第3工序所得的该目标核酸的表达量计算出的乳癌病变部的基因表达量,基于由该判别式获得的结果,预测、判定或评价乳癌患者对曲妥珠单抗发挥治疗敏感性的可能性。

[0055] 本发明进一步在第6方式中,提供上述的任一组合、上述的任一试剂盒、或上述的任一DNA芯片、或它们的组合在用于体外预测、判定或评价乳癌患者对曲妥珠单抗发挥治疗敏感性的可能性的乳癌患者对曲妥珠单抗的治疗敏感性预测用组合物和方法中的应用。

## [0056] 2. 定义

[0057] 本说明书中使用的用语具有以下定义。

[0058] 核苷酸、多核苷酸、利用DNA、RNA等缩略符号的显示,按照“塩基配列又はアミノ酸配列を含む明細書等の制成のためのガイドライン(用于制成包含碱基序列或氨基酸序列的说明书等的指南)”(日本特许厅编)和本技术领域的惯用。

[0059] 本说明书中“多核苷酸”对包含RNA和DNA的任一者的核酸使用。此外,上述DNA包含cDNA、基因组DNA、和合成DNA的任一者。另外上述RNA包含总RNA、mRNA、rRNA、miRNA、siRNA、snoRNA、snRNA、非编码RNA和合成RNA的任一者。另外,在本说明书中,多核苷酸可与核酸互换使用。

[0060] 在本说明书中,“基因”不仅指RNA、和双链DNA,还作为包含构成它们的正链(或正义链)或互补链(或反义链)等的各单链DNA的含义使用。另外对其长度不特别限制。

[0061] 因此,在本说明书中,“基因”只要不特别说明,则包含含有人基因组DNA的双链DNA、含有cDNA的单链DNA(正链)、具有与该正链互补的序列的单链DNA(互补链)、和它们的片段、以及人基因组的任一者。另外该“基因”不仅包含特定的碱基序列(或序列号)所示的“基因”,还包含与由它们所编码的RNA生物学功能等同的RNA、例如编码同族体(即,同源物或直向同源物)、基因多态性等的突变体、和衍生物的“核酸”。作为上述编码同族体、突变体或衍生物的“核酸”,具体可列举具有在后面所记载的严格条件下,与上述序列号1~23所示的任一碱基序列、或该碱基序列中u为t的碱基序列的互补序列杂交的碱基序列的“核酸”。此外,“基因”不考虑功能区的差别,例如可以包含表达控制区、编码区、外显子或内含子。

[0062] 在本说明书中,“转录产物”是指以基因的DNA序列为模板而合成的RNA。RNA聚合酶与位于基因上游的被称为启动子的部位结合,以与DNA的碱基序列互补地使3'末端结合核糖核苷酸下去的形式,合成RNA。该RNA不仅包含基因本身,还包含以表达控制区、编码区、外显子或内含子为代表的从转录起始点到多聚A序列的末端的全序列。

[0063] 另外,本说明书中,“微小RNA(miRNA)”只要不特别说明,则作为包含作为发夹样结构的RNA前体被转录、被具有RNase III切割活性的dsRNA切割酶切割、被称作RISC的蛋白质复合体包入、参与mRNA的翻译抑制的16~25碱基、优选为16~25碱基、更优选为20~25碱基的非编码RNA的含义而使用。另外本说明书中使用的“miRNA”不仅包含特定的碱基序列(或



序列号)所示的“miRNA”,还包含该“miRNA”的前体(pre-miRNA或pri-miRNA),还包含与由它们编码的miRNA生物学功能等同的miRNA、例如编码同族体(即,同源物或直向同源物)、基因多态性等的突变体、和衍生物的“miRNA”。作为上述编码前体、同族体、突变体或衍生物的“miRNA”,具体可列举能够通过miRBase release16 (<http://www.mirbase.org/>) 鉴定的、具有在后面所记载的严格条件下与上述序列号1~23所示的任一特定碱基序列的互补序列杂交的碱基序列的“miRNA”。

[0064] 在本说明书中,“探针”包含用于特异性地检测由基因的表达所生成的RNA或来源于该RNA的多核苷酸的多核苷酸,和/或与其互补的多核苷酸。

[0065] 在本说明书中,“引物”包含特异性地识别、扩增由基因的表达所生成的RNA或来源于该RNA的多核苷酸的连续的多核苷酸和/或与其互补的多核苷酸。

[0066] 这里,互补的多核苷酸(互补链、负链)是指对由用序列号定义的碱基序列、或该碱基序列中u为t的碱基序列组成的多核苷酸的全长序列、或其部分序列(这里为了方便而将其称为正链),基于A:T(U)、G:C这样的碱基对关系而处于碱基上互补的关系的多核苷酸。但是,上述互补链不限于与作为对象的正链的碱基序列完全形成互补序列的情况,只要具有与作为对象的正链在严格条件下能够杂交的程度的互补关系即可。

[0067] 在本说明书中,“严格条件”是指探针与针对其他序列相比,以比能够检测更大的程度(例如背景测定值的平均+背景测定值的标准误差 $\times 2$ 以上的测定值)与其目标序列杂交的条件。严格条件是序列依赖性的,根据进行杂交的环境不同而不同。通过控制杂交和/或洗涤条件的严格性,可以鉴定出对探针为100%互补性的目标序列。

[0068] 在本说明书中,“突变体”,在核酸的情况下,是指由多态性、突变等引起的天然的突变体,或者在序列号1~23所示的碱基序列、或该碱基序列中u为t的碱基序列、或其部分序列中包含1、2或3个以上、优选为1或2个碱基的缺失、替换、添加或插入的突变体,或者在序列号1~23的miRNA的前体RNA的碱基序列、或该碱基序列中u为t的碱基序列、或其部分序列中包含1或2个以上、优选为1或数个碱基的缺失、替换、添加或插入的突变体,或者与上述各个碱基序列或其部分序列显示约90%以上、约95%以上、约97%以上、约98%以上、约99%以上的%同一性的突变体,或者与包含该碱基序列或其部分序列的多核苷酸或寡核苷酸在上述定义严格条件下杂交的核酸。

[0069] 在本说明书中,“数个”是指约20、19、18、17、16、15、14、13、12、11、10、9、8、7、6、5、4、3或2个的整数。

[0070] 在本说明书中,突变体可以使用定点诱变法、利用PCR法的突变引入法等公知的技术制作。

[0071] 在本说明书中,“%同一性”是指将2个序列以最大一致度的方式比对(对齐)时,相对于总碱基数(在有间隙的情况下也包含间隙数)的相同碱基的数的比例(%),可以使用上述的利用BLAST、FASTA的蛋白质/基因的检索系统,在导入间隙或不导入间隙的情况下确定(Karlin,S.等,1993年、Proc.Natl.Acad.Sci.U.S.A.、第90卷、p.5873-5877;Altschul,S.F.等,1990年、Journal of Molecular Biology、第215卷、p.403-410;Pearson,W.R.等,1988年、Proc.Natl.Acad.Sci.U.S.A.、第85卷、p.2444-2448)。

[0072] 在本说明书中,“衍生物”是指修饰核酸,非限定地包含例如,荧光团等的标记化衍生物、包含修饰核苷酸(例如受到了包含卤素、甲基等烷基、甲氧基等烷氧基、硫代、羧甲基

等基团的核苷酸和碱基的再构成、双键的饱和、脱氨基化、氧分子向硫分子的替换等的核苷酸等)的衍生物、PNA(peptide nucleic acid;Nielsen,P.E.et al.,1991,Science254:1497)、LNA(locked nucleic acid;Obika,S.et al.,1998,Tetrahedron Lett.39:5401)等。

[0073] 在本说明书中,“预测、判定、检测或诊断用组合物”是指,为了诊断乳癌的罹患的有无、罹患的程度、乳癌的改善的有无和/或改善的程度、对乳癌的治疗的敏感性,另外为了筛选对乳癌的预防、改善或治疗有用的候选物质,而直接或间接地被利用的组合物。其中包含能够特异性地识别并结合与乳癌的罹患相关地在活体内、特别是乳房组织表达变化的基因的核苷酸、寡核苷酸和多核苷酸。这些核苷酸、寡核苷酸和多核苷酸基于上述性质可以作为用于检测在活体内、组织和/或细胞内等表达的上述基因的探针,另外作为用于扩增在活体内表达的上述基因的引物而有效地利用。

[0074] 在本说明书中,“预测”是指预测、判定、评价、检测或诊断。

[0075] 在本说明书中,成为预测、判定、评价、检测或诊断的对象的“样品”是指随着乳癌的发生、和对乳癌的治疗效果的发挥而本发明的基因发生表达变化的组织和活体材料。具体是指乳房组织及其周围的血管、淋巴结和脏器、以及被怀疑转移的脏器、皮肤、和血液、尿、唾液、汗、组织浸出液等体液、以及粪便、毛发等。

[0076] 本说明书中使用的“FFPE标本”是指将活体组织用福尔马林固定、用石蜡包埋而得的福尔马林固定石蜡包埋标本。

[0077] 本说明书中使用的“对曲妥珠单抗的治疗敏感性”是指通过利用曲妥珠单抗的治疗,乳癌的病情进行被抑制的性质。检测病情的进行的方法可以通过病理学研究,也可以通过利用图像诊断的肿瘤大小的评价和/或患者的状态等临床研究。在对乳癌患者利用曲妥珠单抗的治疗中,可以与曲妥珠单抗一起并用除曲妥珠单抗以外的1种或多种抗癌剂。

[0078] 本说明书中使用的“抗癌剂”是指与曲妥珠单抗并用而在乳癌的药物疗法中使用的药剂。作为抗癌剂,可列举例如,环磷酰胺、塞替派等烷基化剂、氟尿嘧啶、替加氟、卡莫氟、去氧氟尿苷、卡培他滨等5-FU系代谢拮抗剂、甲氨蝶呤、吉西他滨等代谢拮抗药、阿霉素、表柔比星、吡柔比星等蒽环系药、米托蒽醌等蒽醌系药、丝裂霉素C等抗癌抗生素、长春瑞滨等长春花生物碱、紫杉醇、多西他赛等紫杉烷系药、伊立替康等拓扑异构酶I抑制药、他莫昔芬、托瑞米芬等抗雌激素药、法倔唑、阿那曲唑、依西美坦、来曲唑等芳香化酶抑制药、甲羟孕酮等黄体激素剂、戈舍瑞林、亮丙瑞林等LH-RH激动剂、顺铂、卡铂等铂制剂、艾日布林等非紫杉烷系微管动力学抑制剂、拉帕替尼、贝伐单抗、帕妥珠单抗等分子目标药等。

[0079] 本说明书中使用的“顺位”是指在考虑了Rainer,B.等、2004年、FEBS Letters、第573卷、p.83-92所示的假阳性率的统计学的检查中计算的顺位统计量。

[0080] 本说明书中使用的“AUROC值”是指接受者操作特性曲线(ROC曲线)下的面积,成为测量用于将患者分为阳性组和阴性组的预测、判定、评价、检测或诊断的方法的精度的指标。在这些曲线中,关于成为评价对象的方法所显示的结果,将在阳性患者中出现阳性的结果的概率(灵敏度)与在阴性患者中出现阴性的结果的概率(特异性)的倒数作图。

[0081] 在本说明书中“交叉验证(Leave-one-out crossvalidation)法(以下,L00CV法)”是指,将从数据集中除去1个样本而检查对象组和被对象组,制成判别式,用除去的1个样本评价判别式的操作对数据集的全部样本重复而得的评价的平均作为全体的精度的方法。

[0082] 本说明书中使用的“miR-1234基因”或“miR-1234”这样的用语包含序列号1所记载的hsa-miR-1234基因(miRbase Accession No.MIMAT0005589)、该基因的其他生物种同源物或直向同源物等。hsa-miR-1234基因可以通过Lagos-Quintana,M.等,2001年、Science、第294卷、p.853-858所记载的方法得到。

[0083] 本说明书中使用的“miR-513a-5p基因”或“miR-513a-5p”这样的用语包含序列号2所记载的hsa-miR-513a-5p基因(miRbase Accession No.MIMAT0002877)、该基因的其他生物种同源物或直向同源物等。hsa-miR-513a-5p基因可以通过Lagos-Quintana,M.等,2001年、Science、第294卷、p.853-858所记载的方法得到。

[0084] 本说明书中使用的“miR-494基因”或“miR-494”这样的用语包含序列号3所记载的hsa-miR-494基因(miRbase Accession No.MIMAT0002816)、该基因的其他生物种同源物或直向同源物等。hsa-miR-494基因可以通过Lagos-Quintana,M.等,2001年、Science、第294卷、p.853-858所记载的方法得到。

[0085] 本说明书中使用的“miR-26a基因”或“miR-26a”这样的用语包含序列号4所记载的hsa-miR-26a基因(miRbase Accession No.MIMAT0000082)、该基因的其他生物种同源物或直向同源物等。hsa-miR-26a基因可以通过Lagos-Quintana,M.等,2001年、Science、第294卷、p.853-858所记载的方法得到。

[0086] 本说明书中使用的“let-7a基因”或“let-7a”这样的用语包含序列号5所记载的hsa-let-7a基因(miRbase Accession No.MIMAT0000062)、该基因的其他生物种同源物或直向同源物等。hsa-let-7a基因可以通过Lagos-Quintana,M.等,2001年、Science、第294卷、p.853-858所记载的方法得到。

[0087] 本说明书中使用的“let-7b基因”或“let-7b”这样的用语包含序列号6所记载的hsa-let-7b基因(miRbase Accession No.MIMAT0000063)、该基因的其他生物种同源物或直向同源物等。hsa-let-7b基因可以通过Lagos-Quintana,M.等,2001年、Science、第294卷、p.853-858所记载的方法得到。

[0088] 本说明书中使用的“let-7g基因”或“let-7g”这样的用语包含序列号7所记载的hsa-let-7g基因(miRbase Accession No.MIMAT0000414)、该基因的其他生物种同源物或直向同源物等。hsa-let-7g基因可以通过Lagos-Quintana,M.等,2001年、Science、第294卷、p.853-858所记载的方法得到。

[0089] 本说明书中使用的“miR-940基因”或“miR-940”这样的用语包含序列号8所记载的hsa-miR-940基因(miRbase Accession No.MIMAT0004983)、该基因的其他生物种同源物或直向同源物等。hsa-miR-940基因可以通过Lagos-Quintana,M.等,2001年、Science、第294卷、p.853-858所记载的方法得到。

[0090] 本说明书中使用的“miR-1470基因”或“miR-1470”这样的用语包含序列号9所记载的hsa-miR-1470基因(miRbase Accession No.MIMAT0007348)、该基因的其他生物种同源物或直向同源物等。hsa-miR-1470基因可以通过Lagos-Quintana,M.等,2001年、Science、第294卷、p.853-858所记载的方法得到。

[0091] 本说明书中使用的“miR-125a-5p基因”或“miR-125a-5p”这样的用语包含序列号10所记载的hsa-miR-125a-5p基因(miRbase Accession No.MIMAT0000443)、该基因的其他生物种同源物或直向同源物等。hsa-miR-125a-5p基因可以通过Lagos-Quintana,M.等,

2001年、Science、第294卷、p.853-858所记载的方法得到。

[0092] 本说明书中使用的“miR-200c基因”或“miR-200c”这样的用语包含序列号11所记载的hsa-miR-200c基因(miRbase Accession No.MIMAT0000617)、该基因的其他生物种同源物或直向同源物等。hsa-miR-200c基因可以通过Lagos-Quintana,M.等,2001年、Science、第294卷、p.853-858所记载的方法得到。

[0093] 本说明书中使用的“let-7e基因”或“let-7e”这样的用语包含序列号12所记载的hsa-let-7e基因(miRbase Accession No.MIMAT0000066)、该基因的其他生物种同源物或直向同源物等。hsa-let-7e基因可以通过Lagos-Quintana,M.等,2001年、Science、第294卷、p.853-858所记载的方法得到。

[0094] 本说明书中使用的“miR-1228基因”或“miR-1228”这样的用语包含序列号13所记载的hsa-miR-1228基因(miRbase Accession No.MIMAT0005583)、该基因的其他生物种同源物或直向同源物等。hsa-miR-1228基因可以通过Lagos-Quintana,M.等,2001年、Science、第294卷、p.853-858所记载的方法得到。

[0095] 本说明书中使用的“let-7c基因”或“let-7c”这样的用语包含序列号14所记载的hsa-let-7c基因(miRbase Accession No.MIMAT0000064)、该基因的其他生物种同源物或直向同源物等。hsa-let-7c基因可以通过Lagos-Quintana,M.等,2001年、Science、第294卷、p.853-858所记载的方法得到。

[0096] 本说明书中使用的“miR-1229基因”或“miR-1229”这样的用语包含序列号15所记载的hsa-miR-1229基因(miRbase Accession No.MIMAT0005584)、该基因的其他生物种同源物或直向同源物等。hsa-miR-1229基因可以通过Lagos-Quintana,M.等,2001年、Science、第294卷、p.853-858所记载的方法得到。

[0097] 本说明书中使用的“miR-205基因”或“miR-205”这样的用语包含序列号16所记载的hsa-miR-205基因(miRbase Accession No.MIMAT0000266)、该基因的其他生物种同源物或直向同源物等。hsa-miR-205基因可以通过Lagos-Quintana,M.等,2001年、Science、第294卷、p.853-858所记载的方法得到。

[0098] 本说明书中使用的“miR-145基因”或“miR-145”这样的用语包含序列号17所记载的hsa-miR-145基因(miRbase Accession No.MIMAT0000437)、该基因的其他生物种同源物或直向同源物等。hsa-miR-145基因可以通过Lagos-Quintana,M.等,2001年、Science、第294卷、p.853-858所记载的方法得到。

[0099] 本说明书中使用的“miR-181a基因”或“miR-181a”这样的用语包含序列号18所记载的hsa-miR-181a基因(miRbase Accession No.MIMAT0000256)、该基因的其他生物种同源物或直向同源物等。hsa-miR-181a基因可以通过Lagos-Quintana,M.等,2001年、Science、第294卷、p.853-858所记载的方法得到。

[0100] 本说明书中使用的“miR-191基因”或“miR-191”这样的用语包含序列号19所记载的hsa-miR-191基因(miRbase Accession No.MIMAT0000440)、该基因的其他生物种同源物或直向同源物等。hsa-miR-191基因可以通过Lagos-Quintana,M.等,2001年、Science、第294卷、p.853-858所记载的方法得到。

[0101] 本说明书中使用的“miR-125b基因”或“miR-125b”这样的用语包含序列号20所记载的hsa-miR-125b基因(miRbase Accession No.MIMAT0000423)、该基因的其他生物种同

源物或直向同源物等。hsa-miR-125b基因可以通过Lagos-Quintana, M.等, 2001年、Science、第294卷、p.853-858所记载的方法得到。

[0102] 本说明书中使用的“miR-92a基因”或“miR-92a”这样的用语包含序列号21所记载的hsa-miR-92a基因(miRbase Accession No.MIMAT0000092)、该基因的其他生物种同源物或直向同源物等。hsa-miR-92a基因可以通过Lagos-Quintana, M.等, 2001年、Science、第294卷、p.853-858所记载的方法得到。

[0103] 本说明书中使用的“let-7d基因”或“let-7d”这样的用语包含序列号22所记载的hsa-let-7d基因(miRbase Accession No.MIMAT0000065)、该基因的其他生物种同源物或直向同源物等。hsa-let-7d基因可以通过Lagos-Quintana, M.等, 2001年、Science、第294卷、p.853-858所记载的方法得到。

[0104] 本说明书中使用的“miR-23a基因”或“miR-23a”这样的用语包含序列号23所记载的hsa-miR-23a基因(miRbase Accession No.MIMAT0000078)、该基因的其他生物种同源物或直向同源物等。hsa-miR-23a基因可以通过Lagos-Quintana, M.等, 2001年、Science、第294卷、p.853-858所记载的方法得到。

[0105] 本发明提供在乳癌患者对曲妥珠单抗的治疗敏感性的预测(或者,判定、评价、检测或诊断)中有用的组合物、利用了该组合物的乳癌患者对曲妥珠单抗的治疗敏感性的预测(或者,判定、评价、检测或诊断)方法、和利用了该组合物的乳癌患者对曲妥珠单抗的治疗敏感性的预测(或者,判定、评价、检测或诊断)试剂盒,由此,具有提供针对乳癌患者对曲妥珠单抗的治疗敏感性特异性且高预测率、并迅速且简便的预测(或者,判定评价、检测或诊断)方法这样的特别的作用效果。

## 附图说明

[0106] 图1显示用于确定表1所记载的基因的分析的流程。

[0107] 图2显示组合使用与表1所记载的基因对应的序列号1~23的多核苷酸时的乳癌患者对曲妥珠单抗的治疗敏感性的预测率。纵轴显示针对乳癌患者对曲妥珠单抗的治疗敏感性预测的AUROC值,横轴显示在与表1所记载的基因对应的序列号1~23中,对35症例的乳癌患者使用利用LOOCV法的SVM法时的乳癌患者对曲妥珠单抗的治疗敏感性的预测所需的基因的合计数。

[0108] 图3显示以图1所记载的步骤选出的在使用SVM法的乳癌患者对曲妥珠单抗的治疗敏感性预测中使用的20种基因的、利用LOOCV法的选择结果。图3所示的表的行方向显示由34症例的学习数据集和1个测试数据的组合组成的35种教师数据集。列方向显示在各教师数据集中选择的预测用基因的序列号。表中的数字显示在各教师数据集中选择预测用基因时,该基因在第几位被选择的优先顺位。

## 具体实施方式

[0109] 以下进一步具体说明本发明

[0110] 1. 乳癌的目标核酸

[0111] 本发明的上述定义的、作为用于使用乳癌患者对曲妥珠单抗的治疗敏感性预测用组合物和试剂盒来预测乳癌患者对曲妥珠单抗的治疗敏感性的标志物的目标核酸,包含例

如包含序列号1~23所示的碱基序列的人基因(即,分别为miR-1234、miR-513a-5p、miR-494、miR-26a、let-7a、let-7b、let-7g、miR-940、miR-1470、miR-125a-5p、miR-200c、let-7e、miR-1228、let-7c、miR-1229、miR-205、miR-145、miR-181a、miR-191、miR-125b、miR-92a、let-7d、和miR-23a)、它们的同族体、或者它们的突变体或衍生物。这里,基因、同族体、转录产物、突变体和衍生物如上述所定义。优选的目标核酸是包含序列号1~23所示的碱基序列的人基因、它们的转录产物、更优选为该转录产物即miRNA、作为miRNA的前体RNA的pri-miRNA和pre-miRNA。

[0112] 在本发明中,成为用于预测乳癌患者对曲妥珠单抗的治疗敏感性的目标的上述基因,均是具有对曲妥珠单抗的治疗敏感性的乳癌患者相比,在不具有对曲妥珠单抗的治疗敏感性的乳癌患者中,从乳癌病变部得到的基因表达量减少、减低或增加、增大的基因(参照后述的实施例的表1)。

[0113] 第1目标核酸是miR-1234基因、它们的同族体、它们的转录产物、或者它们的突变体或衍生物。到目前为止未知miR-1234基因或其转录产物的表达能够成为乳癌患者对曲妥珠单抗的治疗敏感性的预测标志物这样的报告。

[0114] 第2目标核酸是miR-513a-5p基因、它们的同族体、它们的转录产物、或者它们的突变体或衍生物。到目前为止未知miR-513a-5p基因或其转录产物的表达能够成为乳癌患者对曲妥珠单抗的治疗敏感性的预测标志物这样的报告。

[0115] 第3目标核酸是miR-494基因、它们的同族体、它们的转录产物、或者它们的突变体或衍生物。到目前为止未知miR-494基因或其转录产物的表达能够成为乳癌患者对曲妥珠单抗的治疗敏感性的预测标志物这样的报告。

[0116] 第4目标核酸是miR-26a基因、它们的同族体、它们的转录产物、或者它们的突变体或衍生物。到目前为止未知miR-26a基因或其转录产物的表达能够成为乳癌患者对曲妥珠单抗的治疗敏感性的预测标志物这样的报告。

[0117] 第5目标核酸是let-7a基因、它们的同族体、它们的转录产物、或者它们的突变体或衍生物。到目前为止已知let-7a基因或其转录产物的表达在乳癌患者中表达量降低(上述专利文献1),但未知能够成为乳癌患者对曲妥珠单抗的治疗敏感性的预测标志物这样的报告。

[0118] 第6目标核酸是let-7b基因、它们的同族体、它们的转录产物、或者它们的突变体或衍生物。到目前为止已知let-7b基因或其转录产物的表达在乳癌患者中表达量降低(上述专利文献1),但未知能够成为乳癌患者对曲妥珠单抗的治疗敏感性的预测标志物这样的报告。

[0119] 第7目标核酸是let-7g基因、它们的同族体、它们的转录产物、或者它们的突变体或衍生物。到目前为止未知let-7g基因或其转录产物的表达能够成为乳癌患者对曲妥珠单抗的治疗敏感性的预测标志物这样的报告。

[0120] 第8目标核酸是miR-940基因、它们的同族体、它们的转录产物、或者它们的突变体或衍生物。到目前为止未知miR-940基因或其转录产物的表达能够成为乳癌患者对曲妥珠单抗的治疗敏感性的预测标志物这样的报告。

[0121] 第9目标核酸是miR-1470基因、它们的同族体、它们的转录产物、或者它们的突变体或衍生物。到目前为止未知miR-1470基因或其转录产物的表达能够成为乳癌患者对曲妥

珠单抗的治疗敏感性的预测标志物这样的报告。

[0122] 第10目标核酸是miR-125a-5p基因、它们的同族体、它们的转录产物、或者它们的突变体或衍生物。到目前为止已知miR-125a基因或其转录产物的表达量增加使作为曲妥珠单抗的目标蛋白质的Her2蛋白质的表达量降低(上述非专利文献5),但未知miR-125a的表达量预测Her2阳性乳癌患者对曲妥珠单抗的治疗敏感性这样的报告。

[0123] 第11目标核酸是miR-200c基因、它们的同族体、它们的转录产物、或者它们的突变体或衍生物。到目前为止已知miR-200c基因或其转录产物的表达在乳癌患者中表达量降低(上述专利文献2),但未知能够成为乳癌患者对曲妥珠单抗的治疗敏感性的预测标志物这样的报告。

[0124] 第12目标核酸是let-7e基因、它们的同族体、它们的转录产物、或者它们的突变体或衍生物。到目前为止未知let-7e基因或其转录产物的表达能够成为乳癌患者对曲妥珠单抗的治疗敏感性的预测标志物这样的报告。

[0125] 第13目标核酸是miR-1228基因、它们的同族体、它们的转录产物、或者它们的突变体或衍生物。到目前为止未知miR-1228基因或其转录产物的表达能够成为乳癌患者对曲妥珠单抗的治疗敏感性的预测标志物这样的报告。

[0126] 第14目标核酸是let-7c基因、它们的同族体、它们的转录产物、或者它们的突变体或衍生物。到目前为止未知let-7c基因或其转录产物的表达能够成为乳癌患者对曲妥珠单抗的治疗敏感性的预测标志物这样的报告。

[0127] 第15目标核酸是miR-1229基因、它们的同族体、它们的转录产物、或者它们的突变体或衍生物。到目前为止未知miR-1229基因或其转录产物的表达能够成为乳癌患者对曲妥珠单抗的治疗敏感性的预测标志物这样的报告。

[0128] 第16目标核酸是miR-205基因、它们的同族体、它们的转录产物、或者它们的突变体或衍生物。到目前为止未知miR-205基因或其转录产物的表达能够成为乳癌患者对曲妥珠单抗的治疗敏感性的预测标志物这样的报告。

[0129] 第17目标核酸是miR-145基因、它们的同族体、它们的转录产物、或者它们的突变体或衍生物。到目前为止已知miR-145基因或其转录产物的表达在乳癌患者中表达量降低(上述专利文献3),但未知能够成为乳癌患者对曲妥珠单抗的治疗敏感性的预测标志物这样的报告。

[0130] 第18目标核酸是miR-181a基因、它们的同族体、它们的转录产物、或者它们的突变体或衍生物。到目前为止未知miR-181a基因或其转录产物的表达能够成为乳癌患者对曲妥珠单抗的治疗敏感性的预测标志物这样的报告。

[0131] 第19目标核酸是miR-191基因、它们的同族体、它们的转录产物、或者它们的突变体或衍生物。到目前为止未知miR-191基因或其转录产物的表达能够成为乳癌患者对曲妥珠单抗的治疗敏感性的预测标志物这样的报告。

[0132] 第20目标核酸是miR-125b基因、它们的同族体、它们的转录产物、或者它们的突变体或衍生物。到目前为止已知miR-125b基因或其转录产物的表达量增加使作为曲妥珠单抗的目标蛋白质的Her2蛋白质的表达量降低(上述非专利文献5),但未知miR-125b的表达量预测Her2阳性乳癌患者对曲妥珠单抗的治疗敏感性这样的报告。

[0133] 第21目标核酸是miR-92a基因、它们的同族体、它们的转录产物、或者它们的突变

体或衍生物。到目前为止未知miR-92a基因或其转录产物的表达能够成为乳癌患者对曲妥珠单抗的治疗敏感性的预测标志物这样的报告。

[0134] 第22目标核酸是let-7d基因、它们的同族体、它们的转录产物、或者它们的突变体或衍生物。到目前为止未知let-7d基因或其转录产物的表达能够成为乳癌患者对曲妥珠单抗的治疗敏感性的预测标志物这样的报告。

[0135] 第23目标核酸是miR-23a基因、它们的同族体、它们的转录产物、或者它们的突变体或衍生物。到目前为止未知miR-23a基因或其转录产物的表达能够成为乳癌患者对曲妥珠单抗的治疗敏感性的预测标志物这样的报告。

[0136] 2. 乳癌患者对曲妥珠单抗的治疗敏感性预测用组合物

[0137] 在本发明,可以用于预测乳癌患者对曲妥珠单抗的治疗敏感性的核酸组合物,能够定性和/或定量地测定作为对于乳癌患者对曲妥珠单抗的治疗敏感性的目标核酸的、来自人的miR-1234、miR-513a-5p、miR-494、miR-26a、let-7a、let-7b、let-7g、miR-940、miR-1470、miR-125a-5p、miR-200c、let-7e、miR-1228、let-7c、miR-1229、miR-205、miR-145、miR-181a、miR-191、miR-125b、miR-92a、let-7d、和miR-23a、它们的同族体、或者它们的突变体或衍生物的存在、基因表达量或存在量。

[0138] 上述的目标核酸均与对曲妥珠单抗的治疗敏感性高的乳癌患者相比,在对曲妥珠单抗的敏感性低的乳癌患者中从乳癌组织得到的基因表达量减少、减低或增加、增大。因此,本发明的组合物可以有效地用于对于对曲妥珠单抗的治疗敏感性高的乳癌患者的乳癌组织和对曲妥珠单抗的治疗敏感性低的乳癌患者的乳癌组织,分别测定目标核酸的表达量,将它们进行比较。

[0139] 可以在本发明中使用的组合物,在罹患乳癌的患者样品中,包含:选自包含序列号1~23所示的碱基序列、或该碱基序列中u为t的碱基序列的多核苷酸组及其互补的多核苷酸组、分别与由该碱基序列的互补碱基序列组成的DNA在严格条件下杂交的多核苷酸组及其互补的多核苷酸组、以及包含在这些多核苷酸组的碱基序列中16个以上、优选为21~24个连续的碱基的多核苷酸组中的2种以上的多核苷酸的组合。这些多核苷酸可以作为用于检测作为目标核酸的上述预测用标志物的探针和引物使用。

[0140] 具体地,本发明的组合物可以包含:选自以下所示的多核苷酸、该多核苷酸的突变体、该多核苷酸的衍生物、或该多核苷酸的片段中的2种以上的多核苷酸。

[0141] (1) 由序列号1~23所示的碱基序列、或该碱基序列中u为t的碱基序列组成的多核苷酸、该多核苷酸的突变体、该多核苷酸的衍生物、或该多核苷酸的包含16个以上连续的碱基的片段。

[0142] (2) 包含序列号1~23所示的碱基序列、或该碱基序列中u为t的碱基序列的多核苷酸。

[0143] (3) 由序列号1~9、序列号11~19、和序列号21~23所示的碱基序列、或该碱基序列中u为t的碱基序列组成的多核苷酸、该多核苷酸的突变体、该多核苷酸的衍生物、或该多核苷酸的包含16个以上连续的碱基的片段。

[0144] (4) 包含序列号1~9、序列号11~19、和序列号21~23所示的碱基序列、或该碱基序列中u为t的碱基序列的多核苷酸。

[0145] (5) 由与序列号1~9、序列号11~19、和序列号21~23所示的碱基序列、或该碱基



序列中u为t的碱基序列互补的碱基序列组成的多核苷酸、该多核苷酸的突变体、该多核苷酸的衍生物、或该多核苷酸的包含16个以上连续的碱基的片段。

[0146] (6) 包含与序列号1~9、序列号11~19、和序列号21~23所示的碱基序列、或该碱基序列中u为t的碱基序列互补的碱基序列的多核苷酸。

[0147] (7) 与由在序列号1~9、序列号11~19、和序列号21~23所示的碱基序列中u为t的各个碱基序列的互补碱基序列组成的多核苷酸的任一多核苷酸在严格条件下杂交的多核苷酸、或该多核苷酸的包含16个以上连续的碱基的片段。

[0148] (8) 由序列号10、和序列号20所示的碱基序列、或该碱基序列中u为t的碱基序列组成的多核苷酸、该多核苷酸的突变体、该多核苷酸的衍生物、或该多核苷酸的包含16个以上连续的碱基的片段。

[0149] (9) 包含序列号10、和序列号20所示的碱基序列、或该碱基序列中u为t的碱基序列的多核苷酸。

[0150] (10) 由与序列号10、和序列号20所示的碱基序列、或该碱基序列中u为t的碱基序列互补的碱基序列组成的多核苷酸、该多核苷酸的突变体、该多核苷酸的衍生物、或该多核苷酸的包含16个以上连续的碱基的片段。

[0151] (11) 包含与序列号10、和序列号20所示的碱基序列、或该碱基序列中u为t的碱基序列互补的碱基序列的多核苷酸。

[0152] (12) 与由在序列号10、和序列号20所示的碱基序列中u为t的碱基序列的互补碱基序列组成多核苷酸的任一多核苷酸在严格条件下杂交的多核苷酸、或该多核苷酸的包含16个以上连续的碱基的片段。

[0153] 上述(1)~(12)的多核苷酸的片段可以包含在各多核苷酸的碱基序列、或者各突变体或衍生物的碱基序列中例如连续的16个碱基~序列的总碱基数、16~24、18~24、21~24个碱基等的范围的碱基数,但不限于这些。

[0154] 本发明中使用的上述多核苷酸类或该多核苷酸的片段类均既可以是DNA也可以是RNA。

[0155] 作为本发明的组合物的多核苷酸可以使用DNA重组技术、PCR法、利用DNA/RNA自动合成仪的方法等一般的技术来制作。

[0156] DNA重组技术、定点突变法、PCR法等技术可以使用例如Ausubel等, *Current Protocols in Molecular Biology*, John Willey&Sons, US(1993); Sambrook等, *Molecular Cloning A Laboratory Manual*, Cold Spring Harbor Laboratory Press, US(1989)等所记载的技术。

[0157] 来源于人的miR-1234、miR-513a-5p、miR-494、miR-26a、let-7a、let-7b、let-7g、miR-940、miR-1470、miR-125a-5p、miR-200c、let-7e、miR-1228、let-7c、miR-1229、miR-205、miR-145、miR-181a、miR-191、miR-125b、miR-92a、let-7d、和miR-23a基因是公知,如上所述其取得方法也已知。因此,可以通过克隆这些基因,来制作作为本发明的组合物的多核苷酸。

[0158] 构成本发明的组合物的多核苷酸可以使用DNA自动合成装置来化学合成。该合成一般使用磷酸亚胺法,通过此方法可以自动合成相当于全长的microRNA的单链DNA。DNA自动合成装置由例如Polygen社、Life Technologies社等市售。

[0159] 或者,本发明的多核苷酸还可以通过cDNA克隆化法来制作。cDNA克隆化技术可以利用例如microRNA Cloning Kit Wako(和光纯药工业)等。

[0160] 3. 乳癌患者对曲妥珠单抗的治疗敏感性预测用试剂盒

[0161] 本发明还提供包含与本发明的组合物所含的相同的多核苷酸、该多核苷酸的突变体和/或该多核苷酸的片段的2种以上的乳癌患者对曲妥珠单抗的治疗敏感性预测用试剂盒。

[0162] 本发明的试剂盒包含选自上述2所记载的多核苷酸类的2种以上的多核苷酸、该多核苷酸的突变体、该多核苷酸的衍生物和/或该多核苷酸的片段。这里,突变体和衍生物包含上面定义的含义。

[0163] 本发明的试剂盒可以包含以下物质的2种以上:包含序列号1~9、序列号11~19、和序列号21~23所示的碱基序列、或该碱基序列中u为t的碱基序列的多核苷酸、包含其互补序列的多核苷酸、与这些多核苷酸在严格条件下杂交的多核苷酸、或这些多核苷酸的片段、这些多核苷酸的突变体、或这些多核苷酸的衍生物。

[0164] 本发明的试剂盒进一步可以包含以下物质的1种或2种:包含序列号10和序列号20所示的碱基序列、或该碱基序列中u为t的碱基序列的多核苷酸、包含其互补序列的多核苷酸、与这些多核苷酸在严格条件下杂交的多核苷酸、或这些多核苷酸的片段。

[0165] 可以包含在本发明的试剂盒中的多核苷酸片段例如是选自下述的(1)~(2)中的2种以上的DNA:

[0166] (1) 包含在序列号1~9、序列号11~19、和序列号21~23所示的碱基序列中u为t的碱基序列或其互补序列中16个以上连续的碱基的DNA。

[0167] (2) 除了包含在序列号1~9、序列号11~19、和序列号21~23所示的碱基序列中u为t的碱基序列或其互补的序列中16个以上连续的碱基的DNA之外,进一步包含在序列号10和序列号20所示的碱基序列中u为t的碱基序列或其互补的序列中分别16个以上连续的碱基的DNA。

[0168] 在优选的实施方式中,所述多核苷酸是由序列号1~9、序列号11~19、和序列号21~23的任一项所示的碱基序列、或该碱基序列中u为t的碱基序列组成的多核苷酸、由其互补序列组成的多核苷酸、与这些多核苷酸在严格条件下杂交的多核苷酸、或包含它们的16个以上、优选为21~24个连续的碱基的片段。

[0169] 在其他优选的实施方式中,本发明的试剂盒除了上述的多核苷酸之外,进一步包含由序列号10和序列号20所示的碱基序列、或该碱基序列中u为t的碱基序列组成的多核苷酸、由其互补序列组成的多核苷酸、与这些多核苷酸在严格条件下杂交的多核苷酸、或包含它们的16个以上、优选为21~24个连续的碱基的片段。

[0170] 本发明还提供包含测定由与本发明的组合物所含的相同的多核苷酸、该多核苷酸的突变体和/或该多核苷酸的片段的基因表达量计算的乳癌患者的乳癌组织中的基因表达量的2种以上的乳癌患者对曲妥珠单抗的治疗敏感性预测用试剂盒。

[0171] 在本发明中,上述多核苷酸的片段的大小在上述的各多核苷酸的碱基序列、或者上述的各突变体或衍生物的碱基序列中,例如为连续的16个~序列的总碱基数、16~24、18~24、21~24个碱基等的范围的碱基数。

[0172] 本发明的试剂盒中除了上面说明的本发明中的多核苷酸、该多核苷酸的突变体或

该多核苷酸的片段之外,还可以包含能够预测乳癌患者对曲妥珠单抗的治疗敏感性的已知的或将来发现的多核苷酸。

[0173] 本发明的试剂盒所含的多核苷酸、该多核苷酸的突变体或该多核苷酸的片段可以单独或任意组合地包装在不同的容器中。

[0174] 4. DNA芯片

[0175] 本发明进一步提供包含与本发明的组合物和/或试剂盒所含的相同的多核苷酸(或者,上述的2节的组合物和/或3节的试剂盒所记载的多核苷酸)、突变体、片段和它们的组合的乳癌患者对曲妥珠单抗的治疗敏感性预测用DNA芯片。

[0176] 作为DNA芯片的基板,只要能够固相化DNA就不特别限制,可以例示载玻片、硅制芯片、聚合物制芯片和尼龙膜等。另外这些基板上可以进行聚L赖氨酸涂覆和/或氨基、羧基等的官能团导入等的表面处理。

[0177] 另外对于固相化法只要是一般使用的方法就不特别限制,可以示例:使用称为点样机或点样仪的高密度分注机点样DNA的方法;使用通过压电元件等从喷嘴喷射微少的液滴的装置(喷墨)将DNA吹入基板的方法;或在基板上依次进行核苷酸合成的方法。使用高密度分注机的情况下,例如在具有多个孔的板的各个孔中预先加入不同的基因溶液,将该溶液用针(pin)取出而在基板上依次点样。在喷墨法中,如下进行:从喷嘴喷射基因,在基板上以高速度将基因进行排列配置。在基板上的DNA合成如下进行:用通过光或热脱离的官能团保护结合于基板上的碱基,通过使用掩模来仅对特定部位的碱基施加光或热,使官能团脱离。然后,重复进行在反应液中加入碱基,使其与基板上的碱基偶联的工序。

[0178] 被固相化的多核苷酸是上述所说明的本发明的全部多核苷酸。

[0179] 例如,这样的多核苷酸可以包含选自以下所示的多核苷酸、该多核苷酸的突变体、该多核苷酸的衍生物、或该多核苷酸的片段中的2种以上的多核苷酸。

[0180] (1) 由序列号1~23所示的碱基序列、或该碱基序列中u为t的碱基序列组成的多核苷酸、该多核苷酸的突变体、该多核苷酸的衍生物、或该多核苷酸的包含16个以上连续的碱基的片段。

[0181] (2) 包含序列号1~23所示的碱基序列、或该碱基序列中u为t的碱基序列的多核苷酸。

[0182] (3) 由序列号1~9、序列号11~19、和序列号21~23所示的碱基序列、或该碱基序列中u为t的碱基序列组成的多核苷酸、该多核苷酸的突变体、该多核苷酸的衍生物、或该多核苷酸的包含16个以上连续的碱基的片段。

[0183] (4) 包含序列号1~9、序列号11~19、和序列号21~23所示的碱基序列、或该碱基序列中u为t的碱基序列的多核苷酸。

[0184] (5) 由与序列号1~9、序列号11~19、和序列号21~23所示的碱基序列、或该碱基序列中u为t的碱基序列互补的碱基序列组成的多核苷酸、该多核苷酸的突变体、该多核苷酸的衍生物、或该多核苷酸的包含16个以上连续的碱基的片段。

[0185] (6) 包含与序列号1~9、序列号11~19、和序列号21~23所示的碱基序列、或该碱基序列中u为t的碱基序列互补的碱基序列的多核苷酸。

[0186] (7) 与所述(3)~(6)的任一多核苷酸在严格条件下杂交的多核苷酸、或该多核苷酸的包含16个以上连续的碱基的片段。

[0187] (8) 由序列号10和序列号20所示的碱基序列、或该碱基序列中u为t的碱基序列组成的多核苷酸、该多核苷酸的突变体、该多核苷酸的衍生物、或该多核苷酸的包含16个以上连续的碱基的片段。

[0188] (9) 包含序列号10和序列号20所示的碱基序列、或该碱基序列中u为t的碱基序列的多核苷酸。

[0189] (10) 由与序列号10和序列号20所示的碱基序列、或该碱基序列中u为t的碱基序列互补的碱基序列组成的多核苷酸、该多核苷酸的突变体、该多核苷酸的衍生物、或该多核苷酸的包含16个以上连续的碱基的片段。

[0190] (11) 包含与序列号10和序列号20所示的碱基序列、或该碱基序列中u为t的碱基序列互补的碱基序列的多核苷酸。

[0191] (12) 与所述(8)~(11)的任一多核苷酸在严格条件下杂交的多核苷酸、或该多核苷酸的包含16个以上连续的碱基的片段。

[0192] 在本发明中,上述多核苷酸的片段的大小,对于上述的各多核苷酸的碱基序列、或者上述的各突变体或衍生物的碱基序列,例如为连续的16个~序列的全碱基数、16~24、18~24、21~24个碱基等的范围的碱基数。

[0193] 根据优选的实施方式,本发明的DNA芯片可以包含:包含序列号1~23所示的碱基序列、或该碱基序列中u为t的碱基序列、或该碱基序列的互补序列的多核苷酸的2种以上至全部。

[0194] 在本发明中,被固相化的多核苷酸可以是基因组DNA、cDNA、RNA、合成DNA、合成RNA的任一种,或者既可以是单链也可以是双链。这里,合成DNA和合成RNA包含如上述“衍生物”的定义中所说明的那样的修饰核酸。

[0195] 作为可以检测、测定目标基因、RNA或cDNA的基因表达量的DNA芯片的例子,可以列举东丽株式会社的3D-Gene(注册商标)Human miRNA Oligo chip、Agilent社的Human miRNA Microarray Kit (V2)、EXIQON社的miRCURY LNA(注册商标)microRNA ARRAY等。

[0196] 关于DNA芯片的制作,可以使用例如将预先制备的探针固定化于固相表面的方法。在将预先制备的多核苷酸探针固定化于固相表面的方法中,合成导入了官能团的多核苷酸,在表面处理后的固相载体表面点施寡核苷酸或多核苷酸,使其共价结合(例如,J.B.Lamtur等,Nucleic.Acids.Research、1994年、第22卷、p.2121-2125、Z.Guo等,Nucleic.Acids.Research、1994年、第22卷、p.5456-5465)。多核苷酸一般介由间隔物和/或交联剂与表面处理后的固相载体共价结合。还已知在玻璃表面排列聚丙烯酰胺凝胶的微小片,使合成多核苷酸与其共价结合的方法(G.Yershov等,Proc.Natl.Acad.Sci.U.S.A.、1996年、第94卷、p.4913)。另外,还已知以下方法:在二氧化硅微阵列上制作微小电极的阵列,在电极上设置包含链霉亲和素的琼脂糖的浸透层作为反应部位,通过使该部位带正电荷来固定生物素化多核苷酸,通过控制部位的带电,从而可以实现高速且严密的杂交(R.G.Sosnowski等,Proc.Natl.Acad.Sci.U.S.A.、1997年、第94卷、p.1119-1123)。

[0197] 在利用DNA芯片分析的情况下,使用将本发明的上述诊断用组合物作为DNA探针(单链或双链)贴附于基板而成的DNA芯片。将基因群固相化于基板而得的物体,一般有DNA芯片和DNA阵列这样的名称,DNA芯片包含DNA宏阵列和DNA微阵列,在本说明书中提到DNA芯片时,是指包含该DNA阵列的物体。

### [0198] 5. 乳癌患者对曲妥珠单抗的治疗敏感性预测的检测法

[0199] 本发明提供使用本发明的组合物、试剂盒、DNA芯片、或它们的组合,来体外预测乳癌患者对曲妥珠单抗的治疗敏感性的方法,该方法包括:使用作为手术时或活检检查时采取的样品的乳癌患者的乳癌组织,通过由该诊断用组合物构成的DNA芯片来分析样品中的基因表达量,将显示对曲妥珠单抗的治疗敏感性的乳癌患者的样品中的基因表达量和不显示对曲妥珠单抗的治疗敏感性的乳癌患者的样品中的基因表达量进行比较,在由该样品中的目标核酸的表达量计算的、由乳癌组织获得的基因表达量减少、减低或增加、增大的情况下,预测乳癌患者对曲妥珠单抗的治疗敏感性,其中,该目标核酸可以通过该组合物、试剂盒或DNA芯片中所含的多核苷酸、该多核苷酸的突变体或该多核苷酸的片段来检测。

[0200] 本发明另外提供通过测定本发明的组合物、组合物的基因表达量而构成的试剂盒或DNA芯片的、用于体外预测乳癌患者显示对曲妥珠单抗的治疗敏感性的可能性的用途。

[0201] 在本发明的上述方法中,组合物、试剂盒或DNA芯片使用包含如以上所说明的那样的本发明的多核苷酸、该多核苷酸的突变体或该多核苷酸的片段的单独或者所有可能组合的物质。

[0202] 在本发明的乳癌患者对曲妥珠单抗的治疗敏感性的预测中,本发明的组合物、试剂盒或DNA芯片所含的多核苷酸、该多核苷酸的突变体或该多核苷酸的片段可以作为引物或作为探针使用。在作为引物使用的情况下,可以利用Life Technologies社的TaqMan(注册商标)MicroRNA Assays等,但不限于该方法。

[0203] 本发明的组合物或试剂盒中所含多核苷酸、该多核苷酸的突变体或该多核苷酸的片段,可以在RNA印迹法、DNA印迹法、RT-PCR法、原位杂交法、DNA杂交法等特异性地检测特定基因的公知方法中按照常规方法作为引物或探针利用。作为测定对象样品,根据所使用的检测方法的种类,从作为乳癌患者的手术时或活检检查时采取的样品的乳癌组织的样品中回收。从乳癌患者采取的乳癌组织的样品的状态可以是刚采取时的活的状态,也可以是暂时冻结了的状态,另外也可以是被福尔马林固定了的状态。福尔马林溶液可以使用将市售的福尔马林(甲醛浓度37%)用水稀释而得的,优选使用将用水稀释而得溶液的pH值用碳酸钙、碳酸镁等调整为中性而得的,和/或用磷酸缓冲液稀释而将pH值调整为中性而得的。另外,还可以使用除去恶臭、刺激性气味、调节浓度而得的福尔马林溶液。此外,福尔马林溶液中的甲醛含量优选为1~30%,更优选为2~20%。进而,还可以将被福尔马林固定后的组织被包埋于石蜡中的状态的FFPE标本作为样品。另外进一步可以使用从这些样品/标本按照常规方法制备的总RNA,进而可以使用以该RNA为基础制备的、包含cDNA的各种多核苷酸。

[0204] 样本采取的时期可以是单独的曲妥珠单抗或并用曲妥珠单抗和抗癌剂的治疗开始前也可以是开始后,但优选曲妥珠单抗或并用曲妥珠单抗和抗癌剂的治疗开始前。

[0205] 或者,采取的样品中的本发明的基因、RNA、cDNA等核酸的表达量可以使用DNA芯片进行检测或者定量。此时,本发明的组合物或试剂盒可以作为DNA芯片的探针使用。使所述DNA芯片与以从样品采取的RNA为基础制备的标记DNA或RNA杂交,将该杂交所形成的上述探针与标记DNA或RNA的复合体,以该标记DNA或RNA的标记为指标进行检测,从而可以评价样品中的本发明的乳癌患者对曲妥珠单抗的治疗敏感性预测用组合物的表达的有无或表达的基因的量(基因表达量)。在本发明的方法中,优选使用DNA芯片,这使得对于一个样品同时评价多个基因的表达的有无或基因表达量成为可能。

[0206] 本发明的组合物、试剂盒或DNA芯片在乳癌患者对曲妥珠单抗的治疗敏感性预测中是有用的。具体地,使用该组合物、试剂盒或DNA芯片的乳癌患者对曲妥珠单抗的治疗敏感性预测如下进行:使用作为手术时或活检检查时采取的样品的乳癌患者的乳癌组织,测定样品中的该诊断用组合物的基因表达量,将显示对曲妥珠单抗的治疗敏感性的乳癌患者的样品中的基因表达量、与不显示对曲妥珠单抗的治疗敏感性的乳癌患者的样品中的基因表达量进行比较,从而比较由该样品中的目标核酸的表达量计算出的、由乳癌组织获得的基因表达量减少、减低或增加、增大。此时,基因表达量的不同包含该诊断用组合物的基因表达的有无。

[0207] 利用了本发明的组合物、试剂盒或DNA芯片的乳癌患者对曲妥珠单抗的治疗敏感性的预测方法包括:使用作为活检检查时采取或通过手术摘出乳癌患者的样品的一部分或全部而得的样品的乳癌患者的乳癌组织,使用选自该诊断用组合物的多核苷酸组的一个或多个多核苷酸、该多核苷酸的突变体或该多核苷酸的片段来测定样品中的基因表达量,将显示对曲妥珠单抗的治疗敏感性的乳癌患者的样品中的基因表达量、与不显示对曲妥珠单抗的治疗敏感性的乳癌患者的样品中的基因表达量进行比较,从而比较由该样品中的目标核酸的表达量计算出的、由乳癌组织得到的基因表达量减少、减低或增加、增大,由此预测乳癌患者对曲妥珠单抗的治疗敏感性。

[0208] 本发明的方法例如可以包括以下的工序(a)、(b)和(c):

[0209] 工序(a),使来源于乳癌患者的样品与本发明的组合物、试剂盒或DNA芯片的多核苷酸接触;

[0210] 工序(b),使用上述多核苷酸作为探针来测定样品中的目标核酸的表达量;

[0211] 工序(c),以(b)的结果为基础来预测乳癌患者对曲妥珠单抗的治疗敏感性。

[0212] 作为在本发明方法中使用样品,可列举乳癌患者的样品、例如从乳房组织及其周围组织、疑似乳癌的组织等制备的样品。具体地,由该组织制备的含RNA的样品、或者含有由该样品进一步制备的多核苷酸的样品,可以从活检检查时采取、或通过手术摘出乳癌患者的样品的一部分或全部而得的样品中回收来制备。

[0213] 这里,患者是指罹患乳癌或被强烈怀疑罹患乳癌的哺乳动物,哺乳动物例如非限定地指人、猴、狗、小鼠、大鼠等,优选为人。

[0214] 本发明的方法可以根据作为测定对象使用的样品的种类来变更工序。

[0215] 在作为测定对象物利用RNA的情况下,乳癌患者对曲妥珠单抗的治疗敏感性的预测例如可以包含下述的工序(a)、(b)和(c):

[0216] 工序(a),使由乳癌患者的样品制备的RNA或由该RNA转录的互补多核苷酸(cDNA)与本发明的组合物、试剂盒或DNA芯片的多核苷酸结合;

[0217] 工序(b),使用上述多核苷酸作为探针来测定与该多核苷酸结合的来源于样品的RNA或由该RNA转录的互补多核苷酸;

[0218] 工序(c),基于上述(b)的测定结果来预测乳癌患者显示或不显示对曲妥珠单抗的治疗敏感性。

[0219] 为了通过本发明来预测乳癌患者对曲妥珠单抗的治疗敏感性,可以使用例如各种杂交法。这样的杂交法可以使用例如RNA印迹法、DNA印迹法、PCR法、RT-PCR法、DNA芯片分析法、原位杂交法、DNA杂交法等。

[0220] 在利用RNA印迹法的情况下,可以通过使用本发明的诊断用组合物作为探针,来检测、测定RNA中的各基因表达的有无和/或其基因表达量。具体地可以例示以下方法:将本发明的用于预后预测的诊断用组合物(互补链)用放射性同位素( $^{32}\text{P}$ 、 $^{33}\text{P}$ 、 $^{35}\text{S}$ 等)和/或荧光物质(氰系、罗丹明系、荧光胺系等)等标记,使标记物与按照常规方法转印至尼龙膜等的来源于被测者的样品的RNA杂交,然后用放射线检测器(可以例示BAS-1800II、富士写真フィルム株式会社等)或荧光检测器(可以例示STORM860、GEヘルスケア社等)检测、测定来源于诊断用组合物的标记物(放射性同位素或荧光物质)的信号,从而检测、测定形成的诊断用组合物(DNA)与RNA的双链。

[0221] 在利用定量RT-PCR法的情况下,可以通过使用本发明的上述诊断用组合物中的多核苷酸作为引物,从而检测、测定RNA中的基因表达的有无和/或其基因表达量。具体地可以例示如下方法:从来源于被测者的样品的RNA按照常规方法制备cDNA,以此为模板,使以目标的各基因区能够扩增的方式、以本发明的组合物为基础制备的1对引物(由与上述cDNA结合的正链和负链组成)与cDNA杂交,通过常规方法进行PCR法,从而检测所得的双链DNA。此外,作为双链DNA的检测法,可以采取使用预先用放射性同位素、荧光物质标记好的引物进行上述PCR的方法;将PCR产物用琼脂糖凝胶进行电泳,用溴化乙锭等将双链DNA染色而进行检测的方法;将产生的双链DNA按照常规方法转印至尼龙膜等,并以标记的诊断用组合物中的多核苷酸作为探针与其杂交而进行检测的方法。

[0222] 对杂交条件不限定,例如在 $30^{\circ}\text{C}\sim 60^{\circ}\text{C}$ 、在含有SSC和表面活性剂的溶液中1~24小时的条件。这里,1×SSC是包含150mM氯化钠和15mM柠檬酸钠的水溶液(pH值7.2),表面活性剂包含SDS、Triton、或Tween等。作为杂交条件,更优选包含3~4×SSC、0.1~0.5% SDS。作为杂交后的洗涤条件,可列举例如,通过 $30^{\circ}\text{C}$ 包含0.5×SSC和0.1% SDS的溶液、和 $30^{\circ}\text{C}$ 的包含0.2×SSC和0.1% SDS的溶液、和 $30^{\circ}\text{C}$ 的0.05×SSC溶液的连续的洗涤等条件。互补链优选即使在该条件下洗涤也与作为对象的正链维持杂交状态的互补链。具体地,作为这样的互补链,可以例示由与对象的正链的碱基序列处于完全互补的关系的碱基序列组成的链、以及由与该链具有至少80%、优选为至少85%、更优选为至少90%的同源性的碱基序列组成的链。

[0223] 作为以本发明的组合物或试剂盒的多核苷酸片段作为引物进行PCR时的严格杂交条件的例子,可列举例如,使用由10mM Tris-HCl(pH值8.3)、50mM KCl、1~2mM  $\text{MgCl}_2$ 等组成的PCR缓冲液,在由该引物的序列计算出的 $T_m+5\sim 10^{\circ}\text{C}$ 处理15秒~1分左右等。作为该 $T_m$ 的计算方法,可列举 $T_m=2\times(\text{腺嘌呤残基数}+\text{胸腺嘧啶残基数})+4\times(\text{鸟嘌呤残基数}+\text{胞嘧啶残基数})$ 等。

[0224] 关于这些杂交中的“严格条件”的其他例,例如Sambrook, J. & Russel, D. 著、Molecular Cloning, A LABORATORY MANUAL(分子克隆实验指南)、Cold Spring Harbor Laboratory Press, 2001年1月15日出版、的第1卷7.42~7.45、第2卷8.9~8.17等所记载,可以在本发明中利用。

[0225] 另外,在使用定量RT-PCR法的情况下,可以使用TaqMan商标MicroRNA Assays、Life Technologies社:LNA商标-based MicroRNAPCR、Exiqon社:Ncode商标miRNA qRT-PCT试剂盒、Invitrogen社等为了定量测定miRNA而特别下功夫设计的市售的测定用试剂盒。

[0226] 本发明还提供,使用本发明的组合物、试剂盒、DNA芯片、或它们的组合,测定来源

于乳癌患者的样品中的目标核酸或基因的表达量,利用使用该基因表达量作为教师数据集的SVM法,来预测乳癌患者对曲妥珠单抗的治疗敏感性的方法。

[0227] 即,本发明进一步提供上述方法,包括以下工序:第1工序,使用本发明的组合物、试剂盒、DNA芯片、或它们的组合,体外测定已知显示或不显示对曲妥珠单抗的治疗敏感性的乳癌患者的多个样品中的目标核酸的表达量;第2工序,使用所述第1工序所得的该目标核酸的表达量的测定值,制成利用SVM法预测乳癌患者对曲妥珠单抗的治疗敏感性的判别式;第3工序,与第1工序同样地体外测定由乳癌患者的乳癌组织得到的该目标核酸的表达量;第4工序,在所述第2工序中得到的判别式中将第3工序中得到的该目标核酸的表达量代入判别式,基于由该判别式获得的结果,预测乳癌患者对曲妥珠单抗的治疗敏感性,这里,该目标核酸可以通过该组合物、试剂盒或DNA芯片所含的多核苷酸、该多核苷酸的突变体或该多核苷酸的片段来检测。

[0228] 或者,本发明的方法例如可以包括下述的工序(a)、(b)和(c):

[0229] 工序(a),使用本发明的预测(判定、检测或诊断)用组合物、试剂盒或DNA芯片测定对曲妥珠单抗的治疗敏感性已知的乳癌患者的样品中的目标基因的表达量;

[0230] 工序(b),将(a)所测定的基因表达量的测定值按照下述的步骤代入数2~数5的式中,制成利用SVM法的判别式;

[0231] 工序(c),使用本发明的预测(判定、检测或诊断)用组合物、试剂盒或DNA芯片来测定来源于乳癌患者的样品中的该目标基因的表达量,将它们代入(b)所制成的判别式中,基于所得的结果预测乳癌患者对曲妥珠单抗的治疗敏感性。

[0232] SVM法是1995年AT&T的V.Vapnik在1995年考案的判别分析法(The Nature of Statistical Learning Theory、Springer、1995年)。将应分类的分组已知的数据集的特定数据项目作为解释变量、将应分类的分组作为效标变量,确定用于将该数据集正确地分类于已知的分组的称为超平面的边界面,确定使用该边界面对数据进行分类的判别式。然后,该判别式可以通过将新给予的数据集的测定值作为解释变量代入该判别式,从而预测分组。另外,此时的预测结果既可以是应分类的组,也可以是可以分类于应分类的组的概率,也可以是距超平面的距离(例如,麻生英树等著、統計科学のフロンティア6「パターン認識と学習の統計学新しい概念と手法」(统计科学的边界6“图案识别与学习的统计学新的概念和方法”)、岩波书店(2004年))。

[0233] 本发明的利用SVM法的判别式的解释变量包含通过测定选自上述2节所记载的多核苷酸类中的多核苷酸或该多核苷酸的片段而得到的值,具体地,本发明的用于预测乳癌患者对曲妥珠单抗的治疗敏感性的解释变量是例如选自下述的(1)~(2)中的基因表达量:

[0234] (1)通过包含在序列号1~9、序列号11~19、和序列号21~23所示的碱基序列中u为t的碱基序列或其互补序列中的16个以上连续的碱基的任一DNA测定的乳癌患者的乳癌组织中的基因表达量。

[0235] (2)除了包含在序列号1~9、序列号11~19、和序列号21~23所示的碱基序列中u为t的碱基序列或其互补序列中的16个以上连续的碱基的DNA之外,通过进一步包含在序列号10和序列号20所示的碱基序列中u为t的碱基序列或其互补序列中的分别16个以上连续的碱基的任一DNA测定的乳癌患者的乳癌组织中的基因表达量。

[0236] 以下显示可以在本发明的方法中使用的判别式的计算例。



[0237] 首先,将乳癌患者分为显示对曲妥珠单抗的治疗敏感性的患者组和不显示对曲妥珠单抗的治疗敏感性的患者组的2组。作为判断乳癌患者显示对曲妥珠单抗的治疗敏感性的基准,使用作为显示利用曲妥珠单抗的治疗后的乳癌的病情进行被抑制的状态,在病理学上确认曲妥珠单抗治疗后实施的病理检查结果为被分类为日本乳癌学会编“乳癌取扱い規約第16版”所规定的组织学的治疗效果基准的等级3的病理学的完全奏效、即所有癌细胞陷入坏死或消失的情况、或被肉芽肿样组织或纤维化巢置换的状态作为基准。或者可以以满满足确认所述的组织学的治疗效果基准被分类为等级3的病理学的完全奏效,以及确认临床上没有发现淋巴结转移巢两者为基准。

[0238] 接着,准备由来源于被分开的2组乳癌患者的乳癌组织的活体样品的全面的基因表达量组成的数据集(以下、教师数据集),确定以该2组之间基因表达量可见明确差异的基因作为解释变量、以该分组作为效标变量(例如-1和1)的SVM法的判别式(数1)。此时,该判别式具有数2所定义的约束条件、和数3~5所定义的权重系数(w)和偏置常数(b)。

[0239] 此时的数1~数5由以下的式组成。

[0240] [数1]

$$[0241] \quad f(x) = \sum_{i=1}^n w_i \cdot x_i + b$$

[0242] (这里,x表示由来源于乳癌患者的乳癌组织的活体样品得到的全面的基因表达量组成的数据,x<sub>i</sub>表示由该数据选择出的特定的基因的表达量。)

[0243] [数2]

$$[0244] \quad y_i (w^T x_i + b) \geq 1 - \xi_i$$

$$[0245] \quad \xi_i \geq 0, i = 1, \dots, n$$

[0246] (这里,T表示内积,y表示数据的分类,ξ表示松弛变量。)

[0247] [数3]

$$[0248] \quad \sum \alpha_i - \frac{1}{2} \sum \alpha_i \alpha_j y_i y_j x_i^T x_j$$

[0249] (数3显示使用了通过对数2使用Lagurange(拉格朗日)的乘子法进行回归的Lagurange常数α的最适化问题。)

[0250] [数4]

$$[0251] \quad 0 \leq \alpha_i \leq C, \sum_{i=1}^n \alpha_i y_i = 0$$

[0252] (这里,C表示通过实验确定的约束条件参数。)

[0253] [数5]

$$[0254] \quad w = \sum_{i=1}^n \alpha_i y_i x_i$$

$$[0255] \quad b = -\frac{1}{2} (w^T x_A + w^T x_B)$$

[0256] 而且,对于是否显示对曲妥珠单抗的治疗敏感性未知的乳癌患者,可以通过测定来源于该乳癌患者的乳癌组织的活体样品中的、在该判别式中使用的基因的表达量,将他们代入该判别式的 $x_i$ 中,来预测属于该2组的哪一组。

[0257] 如上所示,在判定是否显示对曲妥珠单抗的治疗敏感性未知的乳癌患者属于显示或不显示对曲妥珠单抗的治疗敏感性的哪一组的判定式的制成中,需要从教师数据集制成的判别式,为了提高该判别式的预测精度,需要在判别式中使用在教师数据集中的2组之间有明确差异的基因。

[0258] 另外,在判别式的解释变量中使用的基因的确定优选如下进行。首先,以作为教师数据集的、来源于显示对曲妥珠单抗的治疗敏感性的乳癌患者组的乳癌组织的活体样品的全面的基因表达量与来源于不显示对曲妥珠单抗的治疗敏感性的乳癌患者组的乳癌组织的活体样品的全面的基因表达量作为数据集,利用作为参数分析的t-检测的p值、作为非参数分析的Mann-Whitney的U检测的p值、或RankProduct法的顺位等,求出该2组间的各基因的表达量的差的大小。

[0259] 接着,使用这里求得的基因表达量的差大的任意的数的基因制成判别式,对该判别式,将来源于其他独立的来源于乳癌患者的乳癌组织的活体样品的基因表达量代入到解释变量中,确定关于该独立的乳癌患者对曲妥珠单抗的治疗敏感性的判别结果。为了构建得到最大的预测精度的判别式,将该判别式的制成和预测精度的确定,一边按照基因表达量的差大的顺序逐一增加基因一边反复评价。

[0260] 另外,为了在该判别式中使用基因的确定和预测精度的确定,优选使用LOOCV法(图1)。即,首先从教师数据集挑出1个数据作为测试数据,将其余作为学习数据集。然后,使用学习数据集制成判别式,使用该判别式预测测试数据所属的组。然后,对能够从教师数据集不重复地分开测试数据的多个组合,更优选对能够不重复地分开测试数据的所有组合确定判别式的预测值,使用确定的预测值和测试数据所属的真的组得到AUROC值,将此作为预测精度。

[0261] 在本发明的方法中,例如,使用基于如上所述的序列号1~23的1个或多个上述多核苷酸、以及/或者基于如上所述的1~23的1个或多个多核苷酸的任意组合,且上述的23种目标基因的表达量全部在显示对曲妥珠单抗的治疗敏感性的乳癌患者与不显示对曲妥珠单抗的治疗敏感性的乳癌患者之间不同,以在来源于乳癌患者的乳癌病变部表达增加/减少为指标,对于这23种基因测定表达量,通过使用这些基因的任意20种基因表达量的组合,从而可以以AUROC值为0.951的预测精度区分乳癌患者对曲妥珠单抗的治疗敏感性(图2)。

[0262] (实施例)

[0263] 通过以下的实施例更具体地说明本发明。但是,本发明的范围不受本实施例限制。

[0264] [实施例1]

[0265] 1. 样品的采取

[0266] 由通过判定Her2蛋白质的免疫组织化学染色法的得分为3+,或所述得分为2+、且荧光原位杂交法的Her2/CEP17比大于2.2而判别为Her2阳性的、获得了知情同意的35症例

的手术前初发乳癌患者,在进行通过曲妥珠单抗与抗癌剂并用的治疗之前使用针活检采取乳癌组织,从采取的乳癌组织得到FFPE标本。而且,从FFPE标本得到切成厚10 $\mu$ m的薄片的乳癌组织的病理标本。

[0267] 具体地,对35症例的Her2阳性乳癌患者在进行通过曲妥珠单抗与抗癌剂的并用的治疗之前进行通过针活检的乳癌组织采取。而且,针活检实施后,通过包含曲妥珠单抗、和环磷酰胺以及多西他赛的术前化学疗法进行治疗。通过曲妥珠单抗和这些抗癌剂的治疗效果的判定基准对于手术时采取的病理检查标本进行,将确认了根据日本乳癌学会编“乳癌取扱い規約第16版”所规定的组织学的治疗效果基准被分类为等级3的病理性完全奏效且临床确认了无淋巴结转移巢的情况,作为有对曲妥珠单抗的治疗敏感性。

[0268] 按照该治疗效果的判定基准,在使用非专利文献3的检查方法的情况下,得知该35症例的Her2阳性乳癌患者中对通过曲妥珠单抗与抗癌剂的并用的治疗显示敏感性的患者数为19症例。即,非专利文献3的检查方法的预测精度为54.2%。

#### [0269] 2. 总RNA的提取

[0270] 从作为样品的以上“1”所得的35症例的Her2阳性乳癌患者的病理标本,使用激光显微切割系统(ライカ社)切下乳癌病变部的组织。由该切下组织,使用Arcturus(注册商标)Paradise(注册商标)Plus2round aminoallyl试剂盒(Life Technologies社),按照该社所规定的步骤得到总RNA。

#### [0271] 3. 基因表达量的测定

[0272] 作为样品使用以上“2”所得的35症例的Her2阳性乳癌患者的总RNA,通过作为寡聚DNA微阵列的3D-Gene(注册商标)Human miRNA Oligo chip(东丽株式会社)来测定基因表达量。寡聚DNA微阵列的测定基于东丽株式会社所规定的步骤来操作,将进行了杂交的DNA微阵列使用3D-Gene(注册商标)扫描仪(东丽株式会社)进行扫描,取得图像并用3D-Gene(注册商标)Extraction(东丽株式会社)将荧光强度数值化。将数值化的荧光强度转换成底为2的对数值作为基因表达量,对35症例的Her2阳性乳癌患者得到通过Human miRNA Oligo chip的各探针杂交而检测的核酸序列、即全面的miRNA的基因表达量。

#### [0273] 4. 预测评分系统

[0274] 将上面“1”~“3”所得的由来源于35症例的Her2阳性乳癌患者的乳癌组织的总RNA检测的miRNA的基因表达量,以上面部分“3”所得的各患者对曲妥珠单抗的治疗敏感性的有无的临床信息为基础在患者间比较,确定对曲妥珠单抗的治疗敏感性预测用基因,使用Matlab version2011a(Mathworks社)计算使用了该基因的情况下的Her2阳性乳癌患者对曲妥珠单抗的治疗敏感性的预测精度。即,按照图1所示的LOOCV法,首先选出上面的部分“3”所得的35症例的Her2阳性乳癌患者组中的75%以上的症例中基因表达量为5以上的miRNA。接着,从35症例的Her2阳性乳癌患者中分出任意的1症例,将该症例的miRNA的基因表达数据作为测试数据。而且,将残余的34症例的miRNA的基因表达数据作为学习数据集。接着,以Her2阳性乳癌患者对曲妥珠单抗的治疗敏感性有无的临床信息作为分组的指标将学习数据集分为2组,对该学习数据集利用RankProduct法进行2组之差的检查,计算出显示学习数据集中的各基因与对曲妥珠单抗的治疗敏感性相关的高度的顺位。接着,使用由RankProduct法得到的顺位最高的1种基因,使用SVM法(数1~数5)制成用于预测对曲妥珠单抗的治疗敏感性的判别式,使用本判别式预测测试数据的对曲妥珠单抗的治疗敏感性。

[0275] 接着,将上述步骤对剩余的34通的所有组合进行,作为结果计算了35通的对曲妥珠单抗的治疗敏感性的预测值。使用该35通的预测值和上面的“1”所得的Her2阳性乳癌患者对曲妥珠单抗的治疗敏感性有无的临床信息计算出的预测精度(AUROC值)为0.540,用35通的组合进行至少1次以上选择而得的基因是序列号1的基因。

[0276] 接下来,为了进一步提高对曲妥珠单抗的治疗敏感性的预测精度,进一步将与对曲妥珠单抗的治疗敏感性的相关高的基因组合。即通过上述的步骤,利用RankProduct法计算顺位,使用第2个以后顺位高的2种以上的基因制成SVM法的判别式,将使用判别式预测测试集对曲妥珠单抗的治疗敏感性的步骤通过LOOCV法对35通的所有组合进行,求出分别的基因数时的预测精度(AUROC值)。

[0277] 其结果是,对曲妥珠单抗的治疗敏感性的预测精度,2基因时AUROC值为0.516、3基因时AUROC值为0.664、4基因时AUROC值为0.714、5基因时AUROC值为0.674、6基因时AUROC值为0.701、7基因时AUROC值为0.707、8基因时AUROC值为0.747、9基因时AUROC值为0.813、10基因时AUROC值为0.816、11基因时AUROC值为0.839、12基因时AUROC值为0.842、13基因时AUROC值为0.780、14基因时AUROC值为0.776、15基因时AUROC值为0.757、16基因时AUROC值为0.707、17基因时AUROC值为0.737、18基因时AUROC值为0.849、19基因时AUROC值为0.901、20基因时AUROC值为0.951、21基因时AUROC值为0.908、并且22基因时AUROC值为0.885,对曲妥珠单抗的治疗敏感性的预测精度在使用20基因时预测精度最大化(图2)。而且,该使用20基因的情况下,用35通的组合进行至少1次以上选择而得的基因为序列号1~23的基因,23种的各基因在35通的组合中被选择的次数如表1。即,显示该使用20基因的本发明的预测精度远远高于非专利文献3的检查方法的预测精度(54.2%)。

[0278] 表1

[0279]

序列号	基因名	在 35 通的LOOCV 法中被选择的次数
1	miR-1234	35
2	miR-513a-5p	35
3	miR-494	35
4	miR-26a	35
5	let-7a	35
6	let-7b	35
7	let-7g	35
8	miR-940	35
9	miR-1470	35
10	miR-125a-5p	35
11	miR-200c	35
12	let-7e	35
13	miR-1228	35
14	let-7c	35
15	miR-1229	35
16	miR-205	33
17	miR-145	33
18	miR-181a	28
19	miR-191	28
20	miR-125b	23
21	miR-92a	20
22	let-7d	8
23	miR-23a	2

[0280] 该AUROC值最大化时,在使用SVM法的乳癌患者对曲妥珠单抗的治疗敏感性预测中使用的20种基因的、通过LOOCV法的筛选结果示于图3。表中的数字表示在各教师数据集中选择预测用基因时该基因在第几位被选择的优先顺位。例如,在35通的各教师数据集中选择与对曲妥珠单抗的治疗敏感性的相关高的1个基因情况下(在图2的曲线中基因数为1的情况下)使用的预测用基因的选择方式有35通,但选择的35通的基因全部为序列号1,在使用序列号1的情况下的对曲妥珠单抗的治疗敏感性的预测精度仅显示AUROC值0.540。另外,

例如,在预测精度(AUROC值)为最大的0.951的从35症例中选出34症例的35通的各教师数据集中,与对曲妥珠单抗的治疗敏感性的相关高的20个基因的组合为序列号1~20;序列号1~19和21;序列号1~19和22;序列号1~19和23;序列号1~18、20和21;序列号1~18、20和22;序列号1~17和19~21;序列号1~17、19、20和22;序列号1~17、19、21和22;序列号1~16和18~21;序列号1~16、18、19、21和22;序列号1~15和17~21;序列号1~15、17~19、22和23的13通的组合。即,在各教师数据集中选择20个基因的情况下(在图2的曲线中基因数为20的情况下)使用预测用基因为序列号1~23的23个基因。

[0281] 在各教师数据集中选择的基因数(图2的曲线中的基因数)为1个~20个的情况下分别使用的预测用基因(序列号)、以及使用该基因时的预测精度(AUROC值)如表2所示。

[0282] 表2

在各教师数据集中被选择的基因数(图2的曲线的基因数)	在各教师数据集中被选择的预测用基因的序列号	预测精度(AUROC值)
1	1	0.540
2	1, 2, 3, 4	0.516
3	1, 2, 3, 4, 7	0.664
4	1, 2, 3, 4, 5, 6, 7	0.714
5	1, 2, 3, 4, 5, 6, 7	0.674
6	1, 2, 3, 4, 5, 6, 7, 8, 9, 10	0.701
7	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11	0.707
8	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14	0.747
9	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14	0.813
10	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14	0.816
[0283] 11	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16	0.839
12	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16	0.842
13	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16	0.780
14	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16	0.776
15	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18	0.757
16	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 21	0.707
17	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21	0.737
18	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22	0.849
19	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23	0.901
20	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23	0.951

[0284] 由以上可知,在乳癌患者对曲妥珠单抗的治疗敏感性预测中,显示最高的AUROC值的基因数是在图2的曲线中基因数为20时,此时使用的预测用基因为表2所示的序列号1~23的23个基因。

[0285] [实施例2]

[0286] 为了确认在使用序列号1~23进行乳癌患者对曲妥珠单抗的治疗敏感性预测的情况下,能够以高于现有方法的精度预测曲妥珠单抗敏感性的最少的基因的组合,对在实施例1中成为对象的35症例的Her2阳性乳癌患者组、和与该患者组不同的48症例的独立的患者组确认了预测精度。

[0287] 1. 从48症例的患者组的样品的采取

[0288] 从通过判定为Her2蛋白质的免疫组织化学染色法的得分为3+、或所述得分为2+且荧光原位杂交法的Her2/CEP17比大于2.2而判别为Her2阳性的、取得了知情同意的与实施例1所收集的症例不同的48症例的手术前初发乳癌患者中,在进行通过曲妥珠单抗与抗癌剂的并用的治疗之前,使用针活检采取乳癌组织,由采取的乳癌组织获得FFPE标本。而且,从FFPE标本切成厚度10 $\mu$ m的薄片而得到乳癌组织的病理标本。

[0289] 具体地,对48症例的Her2阳性乳癌患者在进行通过曲妥珠单抗与抗癌剂的并用的治疗之前进行通过针活检的乳癌组织采取。而且,针活检实施后通过包含曲妥珠单抗、和氟尿嘧啶、表柔比星、环磷酰胺以及多西他赛的术前化学疗法进行治疗。通过曲妥珠单抗与这些抗癌剂的治疗效果的判定基准与实施例1的判定基准相同,对于手术时采取的病理检查标本进行,将确认了依据日本乳癌学会编“乳癌取扱の規約第16版”所规定的组织学的治疗效果的基准被分类为等级3的病理科的完全奏效且临床确认了无淋巴结转移巢的情况作为有对曲妥珠单抗的治疗敏感性。

[0290] 按照该治疗效果的判定基准可知,在使用使用了现在在临床现场被利用的非专利文献3的检查方法的检查法的情况下,该48症例的Her2阳性乳癌患者中,对通过曲妥珠单抗与抗癌剂的并用的治疗显示敏感性的患者数为20症例。即,非专利文献3的检查方法的预测精度为41.7%。

[0291] 2. 从来源于48症例的患者组的样品的总RNA的提取

[0292] 与实施例1的“2”同样地,从作为样品的上面“1”中获得的、48症例的Her2阳性乳癌患者的病理标本中使用激光显微切割系统(ライカ社)切下乳癌病变部的组织。从该切下组织,使用Arcturus(注册商标)Paradise(注册商标)Plus2round aminoallyl试剂盒(Life Technologies社)按照该公司所规定的步骤得到总RNA。

[0293] 3. 来源于48症例的患者组的样品基因表达量的测定

[0294] 与实施例1的“2”同样地,作为样品使用上面“2”所得的48症例的Her2阳性乳癌患者的总RNA,通过作为寡聚DNA微阵列的3D-Gene(注册商标)Human miRNA Oligo chip(东丽株式会社)来测定基因表达量。寡聚DNA微阵列的测定基于东丽株式会社所规定的步骤来操作,将进行了杂交的DNA微阵列使用3D-Gene(注册商标)扫描仪(东丽株式会社)进行扫描,取得图像并用3D-Gene(注册商标)Extraction(东丽株式会社)将荧光强度数值化。将数值化的荧光强度转换为底为2的对数值作为基因表达量,对48症例的Her2阳性乳癌患者,得到通过Human miRNA Oligo chip的各探针杂交而检测的核酸序列、即全面的miRNA的基因表达量。

[0295] 4. 预测评分系统

[0296] 将由实施例1的“1”~“3”所得的来源于35症例的Her2阳性乳癌患者的乳癌组织的总RNA检测的序列号1~23的miRNA的基因表达量,以实施例1的“3”所得的各患者的对曲妥

珠单抗的治疗敏感性的有无的临床信息为基础在患者间进行比较,使用Matlab version2011a(Mathworks社)制成预测使用选自序列号1~23的miRNA的任意2基因时的Her2阳性乳癌患者对曲妥珠单抗的治疗敏感性的预测评分系统。

[0297] 即,从35症例的Her2阳性乳癌患者中分出任意的1症例,将该症例的miRNA的基因表达数据作为测试数据。而且,将剩余的34症例的miRNA的基因表达数据作为学习数据集。接着,以Her2阳性乳癌患者对曲妥珠单抗的治疗敏感性有无的临床信息作为分组的指标将学习数据集分为2组,对该学习数据集,使用SVM法(数1~数5)制成用于使用选自序列号1~23的miRNA的任意2基因来预测对曲妥珠单抗的治疗敏感性的判别式,使用该判别式预测测试数据的对曲妥珠单抗的治疗敏感性。接着,将所述的步骤对其余的34通的全部组合进行,作为结果计算35通的对曲妥珠单抗的治疗敏感性的预测值,求出对35症例的患者组的预测精度(AUROC值)。

[0298] 最终,针对35症例的患者组,对于选自序列号1~23的miRNA的全部2基因的组合求出AUROC值,求出了高于非专利文献3的检查方法的预测精度的65.2%的2基因的组合、及其预测精度。

[0299] 5. 使用了选自序列号1~23的2基因的对曲妥珠单抗的治疗敏感性预测

[0300] 将针对从实施例2的“1”~“3”所得的来源于48症例的Her2阳性乳癌患者的乳癌组织的总RNA检测出的miRNA的基因表达量,使用上面的“4”中制成的、选自序列号1~23的miRNA中的2基因的组合制成的预测评分系统的情况下的Her2阳性乳癌患者对曲妥珠单抗的治疗敏感性的预测精度,使用Matlab version2011a(Mathworks社),对于全部的2基因的组合来确定。

[0301] 其结果是,针对在实施例1中成为对象的35症例的Her2阳性乳癌患者、和在实施例2中成为对象的48症例的Her2阳性乳癌患者的两方对曲妥珠单抗的治疗敏感性的预测精度都变高的、选自实施例1所得的序列号1~23的基因的2基因的组合、及其预测精度如表3所示。即,表3的2基因的组合的本发明的预测精度显示远高于对实施例1的35症例的Her2阳性乳癌患者的非专利文献3的检查方法的预测精度(54.2%)、和对实施例2的48症例的Her2阳性乳癌患者的非专利文献3的检查方法的预测精度(41.7%)的值。

[0302] 表3



[0303]

从 23 个基因中选择 2 个基因的组合		35 症例的 预测结果	48 症例的 预测结果
序列号	序列号	AUROC 值	AUROC 值
4	6	0.671	0.686
4	12	0.658	0.657
5	6	0.678	0.681
5	11	0.655	0.684
5	17	0.678	0.688
5	20	0.668	0.680
5	23	0.681	0.717
6	7	0.671	0.684
6	9	0.674	0.698
6	10	0.655	0.679
6	12	0.664	0.686
6	15	0.678	0.696
6	17	0.664	0.663
6	21	0.671	0.680
6	23	0.694	0.664

[0304] 产业可利用性

[0305] 根据本发明,可以提供预测精度优异的乳癌患者对曲妥珠单抗的治疗敏感性预测用组合物,因而在预测乳癌患者对通过曲妥珠单抗或曲妥珠单抗与抗癌剂的并用的曲妥珠单抗的治疗敏感性中是非常有用的。

	<110> 东丽株式会社 国立大学法人京都大学	
	<120> 乳癌患者对曲妥珠单抗的治疗敏感性预测用组合物和方法	
	<130> PH-5229-PCT	
	<130> JP 2011-097432	
	<130> 2011-04-25	
	<160> 23	
	<170> PatentIn 版本 3.5	
	<210> 1	
	<211> 22	
	<212> RNA	
	<213> 人(Homo sapiens)	
	<400> 1	
[0001]	ueggccugac caccacccc ac	22
	<210> 2	
	<211> 18	
	<212> RNA	
	<213> 人(Homo sapiens)	
	<400> 2	
	uucacagga ggugucau	18
	<210> 3	
	<211> 22	
	<212> RNA	
	<213> 人(Homo sapiens)	
	<400> 3	
	ugaaacauac acggaaacc uc	22
	<210> 4	

	<211> 22	
	<212> RNA	
	<213> 人(Homo sapiens)	
	<400> 4	
	uucaaguaau ccaggauagg cu	22
	<210> 5	
	<211> 22	
	<212> RNA	
	<213> 人(Homo sapiens)	
	<400> 5	
	ugagguagua gguuguauag uu	22
	<210> 6	
	<211> 22	
	<212> RNA	
[0002]	<213> 人(Homo sapiens)	
	<400> 6	
	ugagguagua gguugugugg uu	22
	<210> 7	
	<211> 22	
	<212> RNA	
	<213> 人(Homo sapiens)	
	<400> 7	
	ugagguagua guuuguacag uu	22
	<210> 8	
	<211> 21	
	<212> RNA	
	<213> 人(Homo sapiens)	
	<400> 8	
	aaggcagggc cccgcucce c	21

	<210> 9	
	<211> 21	
	<212> RNA	
	<213> 人(Homo sapiens)	
	<400> 9	
	gcccuccgcc cgugcaccgc g	21
	<210> 10	
	<211> 24	
	<212> RNA	
	<213> 人(Homo sapiens)	
	<400> 10	
	ucccugagac ccuuuaccu guga	24
	<210> 11	
	<211> 23	
	<212> RNA	
[0003]	<213> 人(Homo sapiens)	
	<400> 11	
	uaauacugcc ggguaaugau gga	23
	<210> 12	
	<211> 22	
	<212> RNA	
	<213> 人(Homo sapiens)	
	<400> 12	
	ugagguagga gguuguauag uu	22
	<210> 13	
	<211> 20	
	<212> RNA	
	<213> 人(Homo sapiens)	
	<400> 13	
	ucacaccugc cugcgcgcgc	20

	<210> 14	
	<211> 22	
	<212> RNA	
	<213> 人(Homo sapiens)	
	<400> 14	
	ugagguagua gguuguaugg uu	22
	<210> 15	
	<211> 23	
	<212> RNA	
	<213> 人(Homo sapiens)	
	<400> 15	
	cucucaccac ugcccuicca cag	23
	<210> 16	
	<211> 22	
	<212> RNA	
[0004]	<213> 人(Homo sapiens)	
	<400> 16	
	uccuucauuc caccggaguc ug	22
	<210> 17	
	<211> 23	
	<212> RNA	
	<213> 人(Homo sapiens)	
	<400> 17	
	guccaguuuu cccaggauc ccu	23
	<210> 18	
	<211> 23	
	<212> RNA	
	<213> 人(Homo sapiens)	
	<400> 18	
	aacaucaac gcugucggug agu	23

	<210> 19	
	<211> 23	
	<212> RNA	
	<213> 人(Homo sapiens)	
	<400> 19	
	caacggaauc ccaaaagcag cug	23
	<210> 20	
	<211> 22	
	<212> RNA	
	<213> 人(Homo sapiens)	
	<400> 20	
	ucccugagac cctaacutigu ga	22
	<210> 21	
	<211> 22	
[0005]	<212> RNA	
	<213> 人(Homo sapiens)	
	<400> 21	
	uaaugcacuu gucccggccu gu	22
	<210> 22	
	<211> 22	
	<212> RNA	
	<213> 人(Homo sapiens)	
	<400> 22	
	agagguagua gguugcauag uu	22
	<210> 23	
	<211> 21	
	<212> RNA	
	<213> 人(Homo sapiens)	
	<400> 23	
[0006]	aucacauugc cagggauiuc c	21

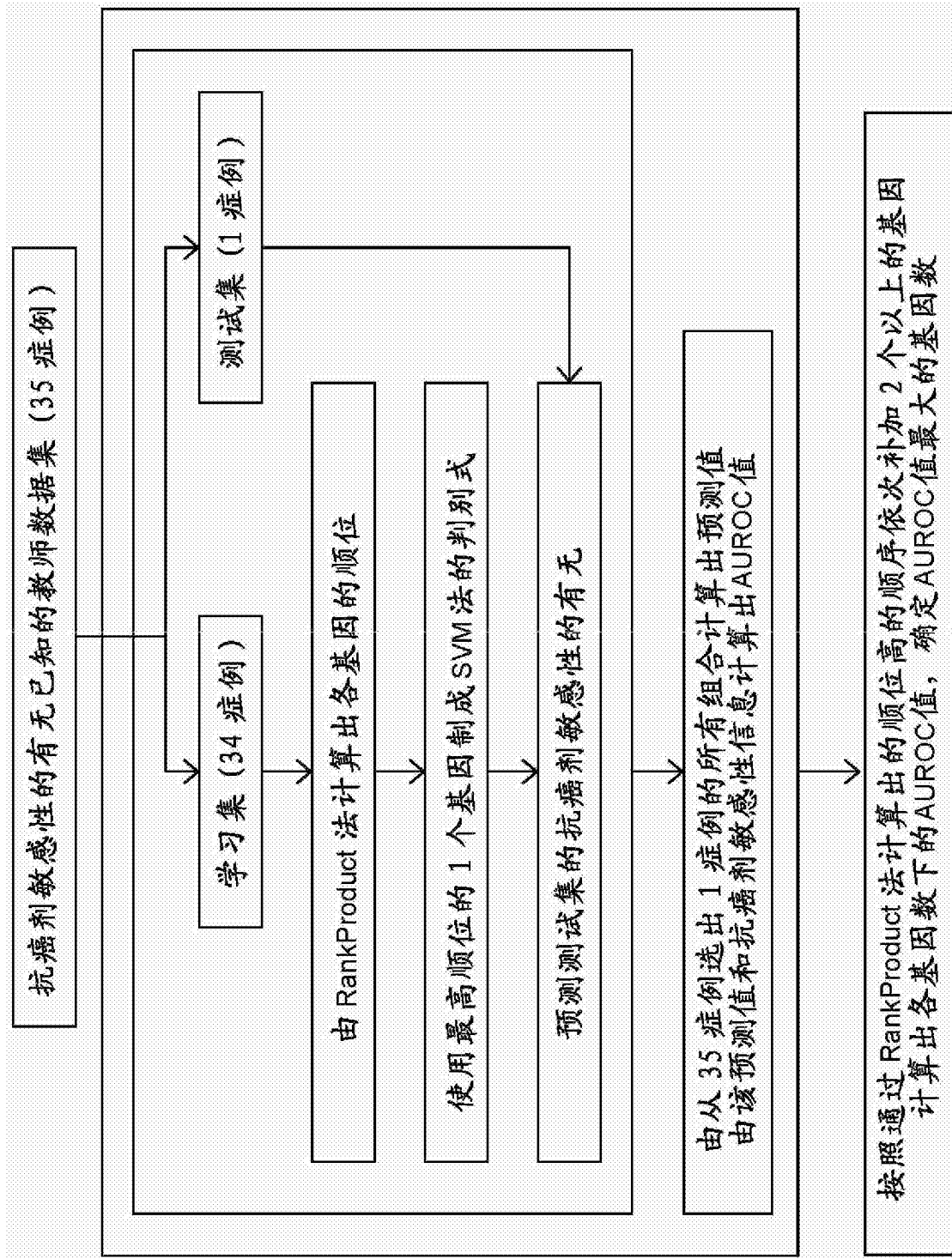


图1

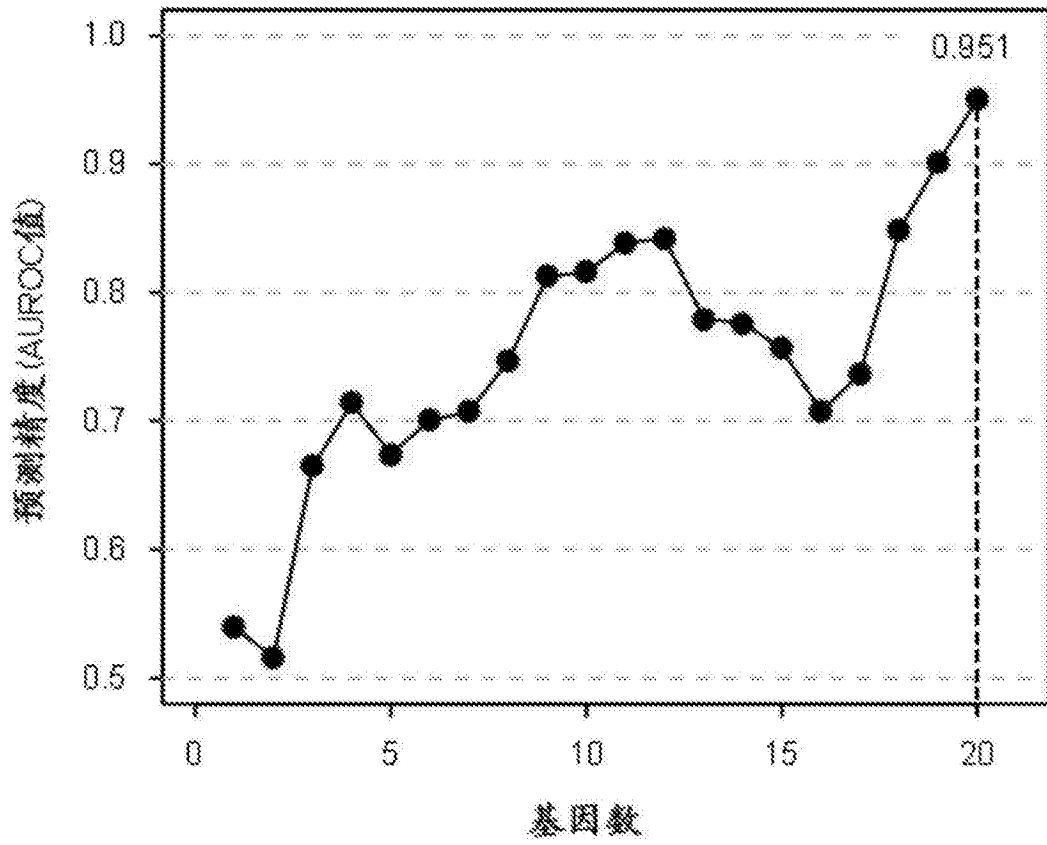


图2



1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12
13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13
14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14
15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15
16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16
17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17
18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18
19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19
20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20
21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21
22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22
23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23

图3