厚生労働科学研究費補助金（創薬基盤推進研究事業）
分担研究報告書

抗癌剤感受性評価系としての移植マウスモデルの確立

分担研究者　岡見　次郎　地方独立行政法人大阪府立病院機構
大阪府立成人病センター　副部長

研究要旨

近年、多くの分子標的薬が開発され、癌薬物治療の選択肢は増えているが、重要なことは個々の患者にとって最も有効な薬剤を予知し、無効な薬剤投与を回避することと考えられる。これまでさまざまなアプローチから癌薬物治療の効果予測に関する研究がなされてきたが、いずれも十分に満足できるレベルに達していない。これまでの研究で、肺癌切除材料を Xenografted マウスに移植し、安定した生着結果を得てきた。そこで本研究の目的は、Xenografted マウスモデルにおける抗癌剤感受性試験により肺癌薬物治療の効果予測のバイオマーカーを探索することである。

今回用いるこの Xenografted マウスモデルでは、第一に個々の患者から得られた腫瘍をマウスに再現し治療効果を評価するため、腫瘍の個体差を考慮に入れた治療効果予測が可能である。次いで、個体ごとの癌細胞を In vitro 培養系を経ずに in vivo で継代するため、培養液（ＦＢＳなど）や培養室（高 CO_2 下培養）による修飾がない。第三に、治療効果を評価するにあたり、画像検査などの間接的な方法によらず、実験動物においてより直接的に測定できる。最後に、本研究では通常困難である治療後のサンプルを含めて解析できる。

60 症例の腫瘍組織を約 125mm^3 に切り分け、免疫不全マウスに移植し患者ごとに Xenografted マウスモデルを作製した。全 60 症例中 27 症例（45%）で移植モデルが確立された。次に、HE 染色を行い、組織学的に検討したところ、腺癌は腺腔構造を有し、また扁平上皮癌は層状構造と部分的な壊死部分を含んでおり、外科切除標本と比べて相同性が観察された。また、免疫組織化学染色による Ki-67 や p53 の発現量の比較でも切除標本と発現量が一致していた。また、培養細胞系と比較して腫瘍組織の組織像や腫瘍細胞の多様性を保持していることが確認された。

これらのマウスモデルを用いた抗癌剤評価が可能であることが確かめられたため、分子標的薬の投与を行い、腫瘍縮小効果及び皮疹関連遺伝子の遺伝子発現を解析する。

A.　研究目的

肺癌は我が国におけるがん死亡原因の第１位（１９％）を占める。手術適応外症例が全体の３分の２を占め、手術例の約半数が血行・リンパ節再発などによって死亡する。これまで術後再発例や非切除例に対してはプラチナ製剤を含む化学療法が行われてきたが、その奏功率は約３０％に過ぎず、一方副作用は必発する。多数の新規抗癌剤が開発され癌薬物治療の選択肢は増えているが、重要なことは個々の患者にとって最も有効な薬剤を予知し、無効な薬剤投与を回避することと考えられる。これまでさまざまなアプローチから癌薬物治療の効果予測に関する研究がなされてきたが、いずれも十分に満足できるレベルに達していない。これまでの研究で、肺癌切除材料をXenograftedマウスに移植し、安定した生着結果を得てきた。そこで本研究の目的は、Xenograftedマウスモデルにおける抗癌剤感受性試験により肺癌薬物治療の効果予測のバイオマーカーを探索することである。

癌薬物治療の効果予測を目的とした研究は、現在まで多く報告されている。しかしその多くは、外科切除標本における分子生物学的な解析と臨床的な治療効果の関連を検討することにより、あるいは樹立され不死化された癌細胞株の腫瘍をマウスで作製し治療効果を解析することによりなされている。今回用いるこのXenograftedマウスモデルは、これらの既存の研究とは異なる新しい方法であり、次のような利点があると考える。第一に個々の患者から得られた腫瘍をマウスに再現し治療効果を評価するため、腫瘍の個体差を考慮に入れた治療効果予測が可能である。同じ疾患の同じ組織型でも癌薬物治療に対する効果は必ずしも一様ではないことから、本研究ではより正確な治療効果予測が期待され、患者さんごとにオーダーメイドな治療法を選択することが可能になる。次いで、個体ごとの癌細胞をIn vitro培養系を経ずにin vivoで継代するため、培養液（ＦＢＳなど）や培養室（高 CO_2 下培養）による修飾がない。実際のヒト腫瘍組織は高度なHeterogeneityを有するので、モノクローナルに増殖する細胞株を用いた動物モデルの解析と比べて、より生体内の腫瘍に近い状態で治療効果を検証することができる。第三に、治療効果を評価するにあたり、画像検査などの間接的な方法によらず、実験動物においてより直接的に測定できる点である。画像検査による腫瘍縮小効果が必ずしも正確に病理学的な殺腫瘍細胞の効果を反映していないことはしばしば経験されるが、本研究では個々の腫瘍の治療効果をより正確に評価することが可能と考えられる。最後に、本研究では通常困難である治療後のサンプルを含めて解析できるという点である。治療前後のサンプルを比較することにより現在までは知られていなかった新たなバイオマーカーが同定される可能性がある。以上より、本研究により今までの臨床サンプルや細胞株を用いた研究では得られなかった新しい知見が得られることが大いに期待される。

B.　研究方法

B－１：移植マウスモデルの確立

2007年1月から2009年3月までに得られた60症例の腫瘍組織を約125mm^{3}に切り分け、免疫不全マウスに移植し患者ごとに

Xenograftedマウスモデルを作製した。増殖が認められた移植片を継代移植するとともに、原発巣の組織との比較をHE染色により行った。

B－２：抗癌剤耐性マウスの発現解析

作製したマウスモデルに対し、継続してシスプラチンの投与を行った。4ヶ月以上投与を続け抵抗性を持つことが確認されたシスプラチン投与、非投与の3症例（6サンプル）に関して発現プロファイル解析を行った。

B－３：倫理面への配慮

本研究では、非小細胞肺癌組織を用いて研究を行う。組織を用いる際は、法令に基づいて設置された大阪府立成人病センター倫理審査委員会において承認を受け、その利用方法を遵守する。また、患者個人に対し主治医もしくはインフォームド・コンセント担当者が説明を行ない、遺伝子解析研究に対する署名、捺印を文書にて頂いている。臨床情報に関しては、大阪府立成人病センター内で匿名化され、その扱いに関しては、センター内の規定に従って扱う。

実験動物に関しては、動物実験等の実施に関する基本指針に基づき設置された大阪府立成人病センター内の動物実験委員会の規定に基づき実験を行う。

C．研究結果

C－１：移植マウスモデルの確立

全60症例中27症例（45%）で移植モデルが確立された。その内訳は、扁平上皮癌が13例（13／24）、小細胞癌が2例（2／2）、多形細胞癌が4例（4／4）、腺癌が8例（8／30）であった。確立した移植マウスは継代可能であり、症例数としては問題ないと考えている。HE染色を行い、組織学的に検討したところ、腺癌は腺腔構造を有し、また扁平上皮癌は層状構造と部分的な壊死部分を含んでおり、外科切除標本と比べて相同性が観察された。

扁平上皮癌の外科切除腫瘍及び移植マウスから得られた腫瘍とではよく似た組織像であった（図1）。

また、肺腺癌の外科切除腫瘍及び移植マウスから得られた腫瘍もよく似た組織像であった（図2）。一方、肺腺癌の培養細胞株PC14をマウスに移植した腫瘍では、外科切除腫瘍とは異なる組織像を示していた（図 2-C）。また、小細胞癌及び転移性大腸癌でも外科切除腫瘍と同様の組織像を示していた（図3, 4）。また、腺癌の症例においては、移植マウスにおける腫瘍で組織内の組織像の多様性が認められた（図 5）。さらに、腎臓及び肝臓への浸潤像も確認された（図6）。

Nude マウスと SCID マウスとの比較では、移植マウスの作製効率に差は認められなかった（表 1）。また、年齢や術後の移植までの時間に関しても作製効率に差は認められなかった（表1）。

次に、免疫組織化学染色による解析を行った。Ki-67、p53 の染色像は外科摘出サンプルと移植マウスとでほぼ一致していた（図7）。また、1サンプルのみの解析ではあるが、CD56、Synapto、Chr-A、PAS に関しても一致していた（表2）。

さらに、シスプラチン耐性を獲得した組織像の観察では、原発巣とのはっきりとした組織的な違いはみられなかった（図8）。

C－２：抗癌剤耐性マウスの発現解析

シスプラチン耐性腫瘍を3症例から確立し、発現プロファイル解析を行った。シスプラチンを投与しない対象群との比較では、P<0.001の23遺伝子を同定した（表3）。

次に、両群で発現に差のある遺伝子群を同定するためにGSEA解析を行った。GSEA解析の結果、S100関連遺伝子が、Enrichment Score(ES)=0.765、Normalized Enrichment Score(NES)=1.935 、 Nominal P-value =0.000、FDR P-value=0.013と有意に偏りがみられた（図9）。また、S100関連遺伝子の2群間のP値はそれぞれ表4に示す通りである。

D．考察・結論

外科切除腫瘍を直接マウスに移植するモデルマウスを確立した。培養細胞系よりも腫瘍組織の組織像や腫瘍細胞の多様性を保持していることが確認された。また、扁平上皮癌に比べ、腺癌の成功率が悪かったが、Nudeマウス及びSCIDマウスともに同じ傾向であり、腫瘍細胞の性質による違いであると考えられる。

これらのマウスモデルを用いた抗癌剤評価が可能であることが確かめられたため（総括研究報告参照）、分子標的薬の投与を行い、腫瘍縮小効果及び皮疹関連遺伝子の遺伝子発現を解析する。また、得られたシスプラチン耐性組織の発現解析を行った。GSEA解析で関連が見出されたS100関連遺伝子のほとんどのS100関連遺伝子が耐性群で高発現になっており、これらの遺伝子がシスプラチンの耐性に関与している可能性が示唆された。

E．研究発表

1．論文発表

(1)「Sublobar resection provides an equivalent survival after lobectomy in elderly patients with early lung cancer.」Okami J, Ito Y, Higashiyama M, Nakayama T, Tokunaga T, Maeda J, Kodama K. Ann Thorac Surg. 2010 Nov;90(5):1651-6.

(2)「Clinical value of F18-fluorodeoxyglucose positron emission tomography-computed tomography in patients with non-small cell lung cancer after potentially curative surgery: experience with 241 patients.」Kanzaki R, Higashiyama M, Maeda J, Okami J, Hosoki T, Hasegawa Y, Takami M, Kodama K. Interact Cardiovasc Thorac Surg. 2010 Jun;10(6):1009-14.

(3)「Prediction of chemotherapeutic effect on postoperative recurrence by in vitro anticancer drug sensitivity testing in non-small cell lung cancer patients.」Higashiyama M, Oda K, Okami J, Maeda J, Kodama K, Imamura F, Minamikawa K, Takano T, Kobayashi H. Lung Cancer. 2010 Jun;68(3):472-7.

2．学会発表

(1)Detection of T790M Gefitinib Resistance Mutation in EGFR using the BEAMing method

Taniguchi K, Nishitani K, Okami J, Kodama K, Higashiyama M, Kato K.

AACR2010 April 19, 2010

(2)Detection of T790M Gefitinib Resistance Mutation in EGFR using the BEAMing method

Taniguchi K, Nishitani K, Okami J, Kodama K, Higashiyama M, Kato K.

第69回日本癌学会総会2010年9月23日 大阪

1. 特許取得

特になし

2. 実用新案登録

特になし

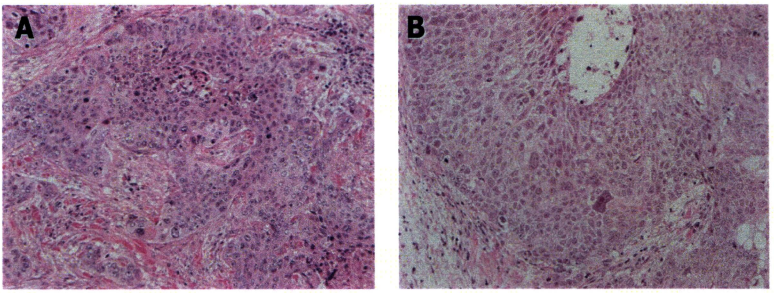


Figure 1. Squamous cell carcinoma histology of surgical (A) and xenograft (B) samples.

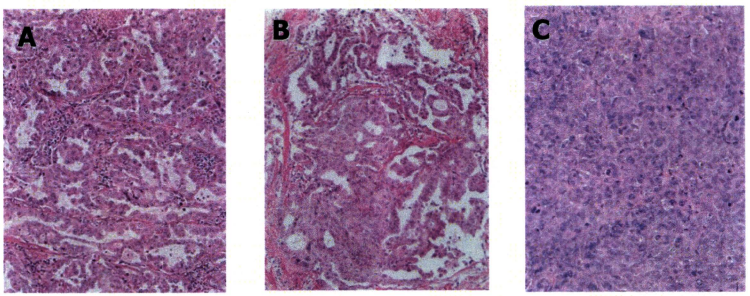


Figure 2. Adenocarcinoma carcinoma histology of surgical (A), xenograft (B), and xenograft tumor of established cell line (PC-14) samples(C).

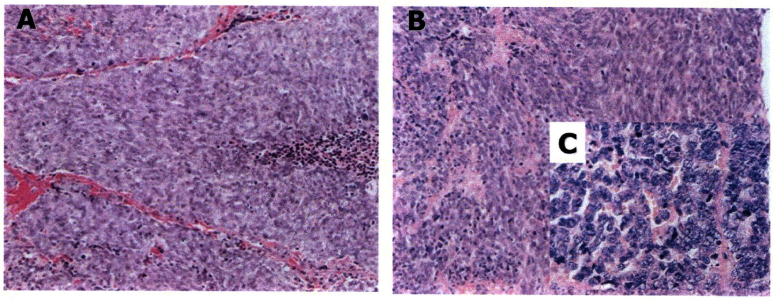


Figure 3. Small cell lung carcinoma histology of surgical (A) and xenograft (B) (C) samples.

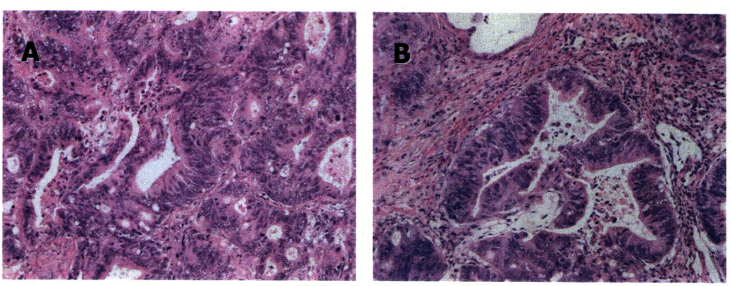
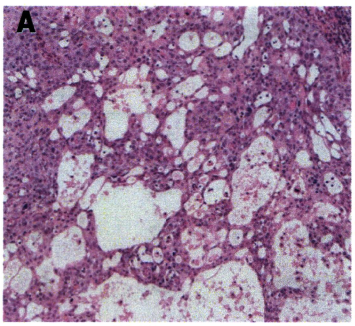


Figure 4. Metastatic colon cancer histology of surgical (A) and xenograft (B) samples.



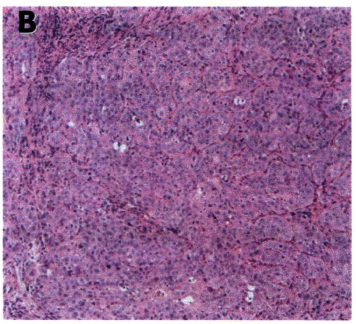
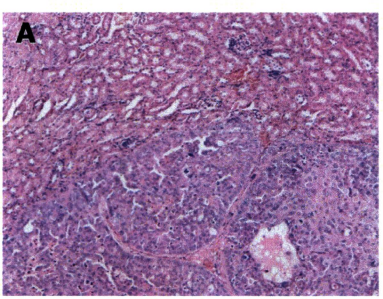


Figure 5. Histological heterogeneity of a xenograft tumor. Histological heterogeneity is often observed in human lung adenocarcinoma. This xenograft tumor derived from lung adenocarcinoma showed glandular pattern (A) and less differentiated cobble stone pattern (B).



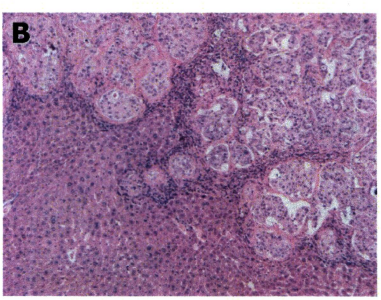


Figure 6. Direct invasion into adjacent organs of xenograft tumors. Invasion is one of the hall-marks of malignant tumor. When cancer cells were implanted into peri-renal pocket (backside of abdominal cavity), xenograft tumor grew in abdominal cavity and invaded into adjacent organ including kidney (A), liver (B), and ovary.

Table 1. Tumorigenecity and experimental factors

Experimental factors		Fail	Success
Type of mouse	Nude	11	7
	SCID	12	13
Age of mouse (weeks old)	4 or 5 w.o.	12	14
	6 w.o. or older	11	6
Timing of implanting	Within 3 hours	15	12
(Hours after surgery)	Longer than 3 hours	8	8

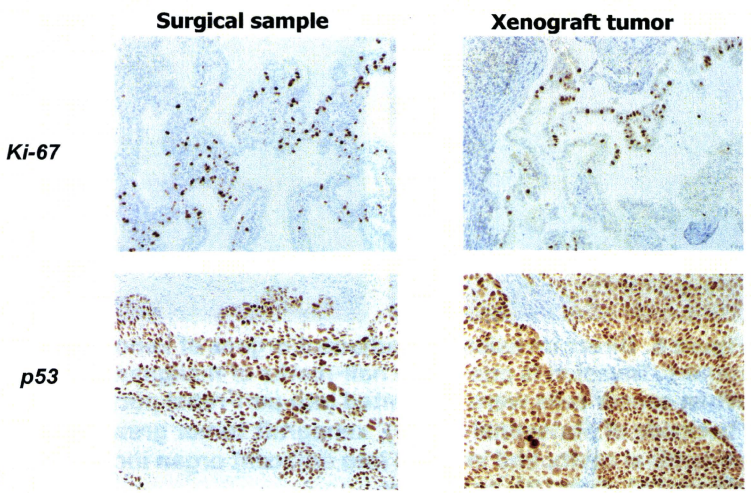


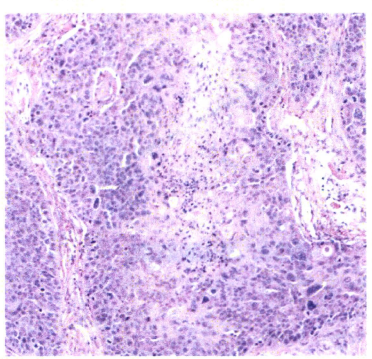
Figure7.Immunohistochemical and mutational analysis

Table2.Immunohistochemical and mutational analysis

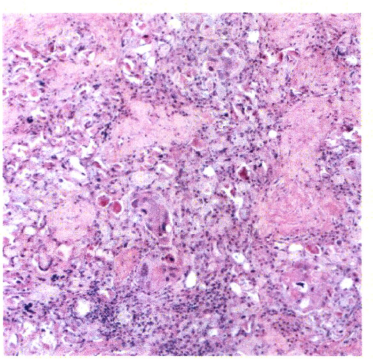
Sample ID	Histology	Ki-67		p53		CD56		Synapto		Chr-A		PAS	
		S	X	S	X	S	X	S	X	S	X	S	X
441921	Small	++	++	-	-	+	+	-	-	+	+		
465111	Small	+++	NE	+	+	+	+	-	-	Focal	-		
460011	Squamous	++	++	-	-								
477711	Squamous	+++	+++	+	+								
487711	Adenoca	+++	+++									Focal	Focal
499811	Adenoca	+	+	-	-								

S: Surgical sample, X; Xenografted tumor
All samples were negetive for k-ras or EGFR mutation.
Semiquantitative analysis of ki-67: + ~30%, ++ 31~60%, +++ 61%~

Figure8.Pathological evaluation of chemo-response



Control



After CDDP treatment

表 3：シスプラチン耐性組織で発現が変動した遺伝子リスト

NO.	Gene Name	Fold change	p-value
A_23_P334263	SUMO/sentrin specific peptidase family member 8	0.27	1.25E-05
A_24_P89891	TNF receptor-associated factor 1	1.60	0.000104
A_24_P30923	stannin	-0.35	0.000123
A_23_P26004		0.58	0.000148
A_32_P113887		-0.43	0.000269
A_24_P380734	syndecan 2	0.76	0.000308
A_23_P200930	5-methyltetrahydrofolate-homocysteine methyltransferase	0.20	0.00037
A_24_P33197	RANBP2-like and GRIP domain containing 2	0.90	0.000384
A_32_P121908		0.73	0.000415
A_24_P109069	synaptotagmin XV	0.54	0.000431
A_23_P110022	GATA binding protein 2	-0.73	0.000563
A_24_P306479		2.38	0.000609
A_23_P11752	opioid receptor, delta 1	-0.70	0.000633
A_32_P86820		1.19	0.000695
A_23_P418485	chromosome 11 open reading frame 65	0.93	0.000699
A_23_P139786	2'-5'-oligoadenylate synthetase-like	-1.89	0.000751
A_23_P171034	NAD(P) dependent steroid dehydrogenase-like	-0.51	0.000773
A_24_P520241		0.43	0.000828
A_23_P120227	limb bud and heart development homolog (mouse)	0.90	0.000854
A_23_P428738	angiogenin, ribonuclease, RNase A family, 5	1.75	0.0009
A_23_P427075	cystinosis, nephropathic	-0.59	0.00096
A_23_P58132	ras homolog gene family, member H	0.19	0.00096
A_24_P323698		0.36	0.000969

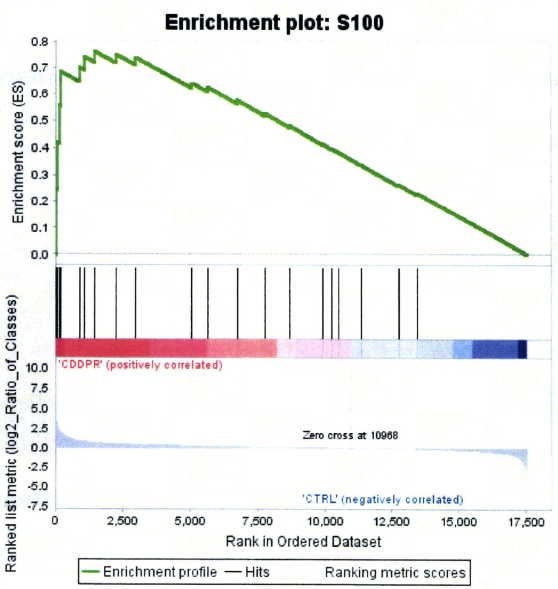


図 9：S100 関連遺伝子の GSEA 解析結果

表 4：S100 関連遺伝子のシスプラチン耐性組織における遺伝子発現

NO.	Gene Name	Fold change	p-value
A_23_P383227	S100 calcium binding protein A1	-0.09	0.845505
A_23_P137984	S100 calcium binding protein A10	0.12	0.597834
A_23_P126593	S100 calcium binding protein A11	-0.35	0.671698
A_23_P145863	S100 calcium binding protein A11	-0.33	0.592599
A_23_P74001	S100 calcium binding protein A12	-3.88	0.018828
A_23_P372874	S100 calcium binding protein A13	-0.36	0.236035
A_23_P124619	S100 calcium binding protein A14	-1.20	0.143129
A_23_P147918	S100 calcium binding protein A16	-1.01	0.139555
A_23_P201706	S100 calcium binding protein A2	-0.16	0.911855
A_23_P104073	S100 calcium binding protein A3	-0.51	0.948266
A_23_P94800	S100 calcium binding protein A4	-0.90	0.037741
A_23_P115467	S100 calcium binding protein A5	-0.95	0.637009
A_23_P201711	S100 calcium binding protein A6	-0.39	0.470088
A_23_P103310	S100 calcium binding protein A7	-2.67	0.015943
A_24_P280274	S100 calcium binding protein A7A	-1.12	0.545125
A_23_P434809	S100 calcium binding protein A8	-3.34	0.015398
A_23_P23048	S100 calcium binding protein A9	-4.37	0.04339
A_23_P143526	S100 calcium binding protein B	-0.32	0.917114
A_23_P44996	S100 calcium binding protein G	0.19	0.134793
A_23_P58266	S100 calcium binding protein P	-1.59	0.175542
A_24_P12136	S100 calcium binding protein Z	0.13	0.195099

Ⅲ. 研究成果の一覧　雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Kawarazaki S, Taniguchi K, Shirahata M, Kukita Y, Kanemoto M, Mikuni N, Hashimoto N, Miyamoto S, Takahashi JA, Kato K.	Conversion of a molecular classifier obtained by gene expression profiling into a classifier based on real-time PCR: a prognosis predictor for gliomas.	BMC Med Genomics.	Nov 10;	3:52.	2010
Taniguchi K, Yamada T, Sasaki Y, Kato K.	Genetic and epigenetic characteristics of human multiple hepatocellular carcinoma.	BMC Cancer.	Oct 6;	10:530.	2010
Okami J, Ito Y, Higashiyama M, Nakayama T, Tokunaga T, Maeda J, Kodama K.	Sublobar resection provides an equivalent survival after lobectomy in elderly patients with early lung cancer.	Ann Thorac Surg.	Nov;90(5):	1651-6.	2010
Kanzaki R, Higashiyama M, Maeda J, Okami J, Hosoki T, Hasegawa Y, Takami M, Kodama K.	Clinical value of F18-fluorodeoxyglucose positron emission tomography-computed tomography in patients with non-small cell lung cancer after potentially curative surgery: experience with 241 patients.	Interact Cardiovasc Thorac Surg.	Jun;10(6):	1009-14.	2010
Higashiyama M, Oda K, Okami J, Maeda J, Kodama K, Imamura F, Minamikawa K, Takano T, Kobayashi H.	Prediction of chemotherapeutic effect on postoperative recurrence by in vitro anticancer drug sensitivity testing in non-small cell lung cancer patients.	Lung Cancer.	Jun;68(3):	472-7.	2010

Kawarazaki et al. BMC Medical Genomics 2010, 3:52
http://www.biomedcentral.com/1755-8794/3/52

BMC Medical Genomics

TECHNICAL ADVANCE Open Access

Conversion of a molecular classifier obtained by gene expression profiling into a classifier based on real-time PCR: a prognosis predictor for gliomas

Satoru Kawarazaki[1,2], Kazuya Taniguchi[1], Mitsuaki Shirahata[2], Yoji Kukita[1], Manabu Kanemoto[1,2], Nobuhiro Mikuni[2], Nobuo Hashimoto[3], Susumu Miyamoto[2], Jun A Takahashi[4], Kikuya Kato[1*]

Abstract

Background: The advent of gene expression profiling was expected to dramatically improve cancer diagnosis. However, despite intensive efforts and several successful examples, the development of profile-based diagnostic systems remains a difficult task. In the present work, we established a method to convert molecular classifiers based on adaptor-tagged competitive PCR (ATAC-PCR) (with a data format that is similar to that of microarrays) into classifiers based on real-time PCR.

Methods: Previously, we constructed a prognosis predictor for glioma using gene expression data obtained by ATAC-PCR, a high-throughput reverse-transcription PCR technique. The analysis of gene expression data obtained by ATAC-PCR is similar to the analysis of data from two-colour microarrays. The prognosis predictor was a linear classifier based on the first principal component (PC1) score, a weighted summation of the expression values of 58 genes. In the present study, we employed the delta-delta Ct method for measurement by real-time PCR. The predictor was converted to a Ct value-based predictor using linear regression.

Results: We selected *UBL5* as the reference gene from the group of genes with expression patterns that were most similar to the median expression level from the previous profiling study. The number of diagnostic genes was reduced to 27 without affecting the performance of the prognosis predictor. PC1 scores calculated from the data obtained by real-time PCR showed a high linear correlation ($r = 0.94$) with those obtained by ATAC-PCR. The correlation for individual gene expression patterns ($r = 0.43$ to 0.91) was smaller than for PC1 scores, suggesting that errors of measurement were likely cancelled out during the weighted summation of the expression values. The classification of a test set ($n = 36$) by the new predictor was more accurate than histopathological diagnosis (log rank p-values, 0.023 and 0.137, respectively) for predicting prognosis.

Conclusion: We successfully converted a molecular classifier obtained by ATAC-PCR into a Ct value-based predictor. Our conversion procedure should also be applicable to linear classifiers obtained from microarray data. Because errors in measurement are likely to be cancelled out during the calculation, the conversion of individual gene expression is not an appropriate procedure. The predictor for gliomas is still in the preliminary stages of development and needs analytical clinical validation and clinical utility studies.

* Correspondence: katou-ki@mc.pref.osaka.jp
[1]Research Institute, Osaka Medical Center for Cancer and Cardiovascular Diseases, 1-3-3 Nakamichi, Higashinari-ku, Osaka, 537-8511, Japan
Full list of author information is available at the end of the article

Background

Since the inception of gene expression profiling, researchers have sought to use this technology to improve the diagnosis of diseases, especially cancers. Recently, MammaPrint [1,2] and Oncotype DX [3,4] were established as diagnostic tests based on multiple gene assays for breast cancer. Despite the success of these diagnostic tests, the development of assays for gene expression profiling is still difficult. In particular, there have been few examples of microarray-based diagnostic tests, although microarrays are frequently used as a discovery tool. One reason for the paucity of microarray-based diagnostic tests is that DNA microarrays require considerable effort to achieve the level of technical refinement necessary for diagnostic practice. On the contrary, real-time PCR is stable and robust and is frequently used for diagnosis. Because there are many studies describing the use of microarrays at the discovery phase, a convenient method to convert a microarray-based algorithm into one based on real-time PCR would help to accelerate the development of diagnostic systems based on gene expression profiling.

Previously, we performed gene expression profiling of 152 glioma tissues [5] with a high-throughput quantitative PCR technique called adaptor-tagged competitive PCR (ATAC-PCR) [6,7]. ATAC-PCR is an advanced version of quantitative competitive PCR characterised by the addition of unique adaptors for different cDNAs. A single ATAC-PCR reaction includes five cDNA samples and two different amounts of a control cDNA sample with different adaptor tags, and it measures the relative expression of the samples against that of the control. We discovered a correlation between gene expression profiles and glioma prognosis, and we developed a prognosis predictor based on a 58-gene profile [5]. The performance of the predictor based on ATAC-PCR was cross-validated with a learning set of 110 glioma samples and validated with a test set of 42 samples. Cox regression analysis revealed that the correlation between the predictor and the prognosis was superior to that of histological classification and was an independent risk factor. The current prognostic standard, the histopathological classification system, is limited in its diagnostic accuracy, and prognoses range widely even within the same grade. Diagnosis depends on individual pathologists, and the results are often discordant among multiple pathologists [8]. The performance of the prognosis predictor based on ATAC-PCR indicated that this predictor held promise for the support of conventional histopathological classification. Our classifier is also expected to bring benefits in the clinical setting for personalized management of glioma patients. For example, various molecular-targeted drugs have recently been evaluated in clinical trials for gliomas. These novel treatments should be considered for tumours that are resistant to conventional chemoradiotherapy. Yet, it is important to avoid using such a therapy for tumours that are sensitive to conventional chemoradiotherapy, based on the cost and adverse effects associated with this technique. Considering elevated expression of angiogenesis-related genes in the poor prognosis group, [5], our classifier might be useful for selection of patients for anti-VGEF agents.

In the present study, we converted the conventional predictor to one based on real-time PCR. This new predictor is based on the delta-delta Ct method [9] and requires only the measurement of the cycle threshold (Ct) of diagnostic genes. For the conversion, we first identified a reference gene for real-time PCR. Then we constructed the parameters for the conversion formula using data obtained from the learning set, which was used to construct the original classifier. Finally, the new classifier was validated with a test set. Because there is a linear correlation between microarray data and Ct values [10], the conversion process could be applicable for classifiers based on microarrays.

Methods

Patients and tumour samples

Specimens excised from 80 patients with high-grade glioma (69 cases of glioblastoma and 11 cases of anaplastic astrocytoma) at Kyoto University Hospital or nearby regional hospitals between 1998 and 2008 were stored at -70°C until use. All histological diagnoses were performed in the Kyoto University Pathology Unit according to the 2000 or 2007 WHO classifications.

Sixty of the 80 samples were recruited from those used in the previous study [5]. They were collected from patients enrolled in a phase II clinical trial using nimustine, carboplatin, vincristine, and IFN-β with radiotherapy for high-grade gliomas (the KNOG study) [11]. The remaining 20 patients were treated with temozolomide and radiotherapy. The learning set included 44 samples (43 glioblastoma, 1 anaplastic astrocytoma) from the KNOG study. Recurrence was detected in 36 of the 44 patients and their median progression-free survival was 7 months. The test set included 36 samples (26 glioblastoma and 10 astrocytoma). Twenty-three of the 36 patients showed tumour progression, and their median progression-free survival was 8 months.

Institutional approval for this study was obtained from the Institutional Review Board of Kyoto University, and informed consent was obtained from all patients prior to surgery.

RNA extraction and cDNA synthesis

Total RNA was isolated from 100 mg of the tumour specimen using TRIzol (Invitrogen, Carlsbad, CA, USA)

according to the manufacturer's instructions. RNA concentrations and A260/A280 ratios were measured using a NanoDrop ND-1000 (NanoDrop Technologies, Montchanin, DE, USA). Only RNA samples with A260/A280 ratios above 1.90 were included in the study. RNA integrity was confirmed by analysis with the Agilent 2100 bioanalyser.

After DNase treatment, 5 μg of total RNA in 10 μl of distilled water was incubated with 1 μl of oligo(dT) primer for 5 min at 70°C. Total RNA was reverse transcribed in a total volume of 20 μl containing 4 μl of 5× first strand buffer, 1 μl of RNase inhibitor (Invitrogen), 2 μl of 0.1 M DTT, 0.5 μl of 20 mM dNTP and 1 μl of SuperScript III Reverse Transcriptase (Invitrogen). The samples were incubated at 45°C for 1 hr. Next, a reaction mixture (total volume of 103 μl) containing 10 μl of 10× *Escherichia coli* (*E. coli*) ligation buffer, 2 μl of 20 mM dNTPs, 2 μl of 0.1 M DTT, 2 μl of *E. coli* ligase (Invitrogen), 1 μl of RNase H (Invitrogen), 4 μl of *E. coli* DNA polymerase (Invitrogen) and 82 μl of nuclease-free water was added. The resulting reaction mixture was incubated at 16°C for 120 min and then at 70°C for 20 min. The reaction mixture was then diluted five-fold with nuclease-free water and stored at -30°C until RT-PCR analysis.

Primer design and optimisation

Gene sequences were retrieved using the UCSC Genome Bioinformatics http://genome.ucsc.edu/ program, and primers sequences were designed using Primer3Plus http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi. Specific interactions between primers and target genes were confirmed using either NCBI BLAST http://blast.ncbi.nlm.nih.gov/Blast.cgi) or BlastView (http://uswest.ensembl.org/index.html. The specificity of the expected RT-PCR products was determined based on melting curve analyses of reactions with glioma cDNA and human cDNA libraries. The product-specific melting curves showed only single peaks and no primer-dimer peaks or artefacts.

Quantitative real-time reverse transcription-PCR

Quantitative PCR amplification assays were performed by a SYBR Green fluorescent assay using the ABI PRISM 7500 real-time PCR sequence detection system (Applied Biosystems, Foster City, CA, USA). Reactions were performed in a 96-well plate with 20-μl reaction solutions containing SYBR *Premix Ex Taq* II (10 μl) (Takara Bio., Inc., Japan), ROX reference dye II (0.4 μl), 10 μM forward and reverse primers (0.8 μl), 1 μl of cDNA template, and nuclease-free water (7 μl). Cycling conditions included an initial denaturation for 10 sec at 95°C, followed by 40 cycles of 5 sec at 95°C and 34 sec at 60°C. For determination of the reference gene, a standard curve was generated for each assay using seven serial dilutions of an amplified human brain cDNA library ranging from 20 ng to 20 fg.

The delta-delta Ct method was employed for the diagnostic assays. Ct values were calculated following the manufacturer's instructions (Applied Biosystems, Foster City, CA, USA), using *UBL5* as the internal reference. The diagnostic genes fulfilled the criterion that the absolute value of the slope of the log input amount vs. ΔCt should be < 0.1.

Data analysis

Thirty primers for the selected gene candidates and for the internal and negative controls were added in triplicate to 96-well plates, and the samples were measured using one plate per sample. The negative controls showed no detectable amplification or background levels of amplification (Ct ≥ 37, compared with 16 to 31 with sample DNAs). The mean and the standard deviation of differences of Ct values between duplicates were 0.060 and 0.086, respectively. Sequence detection software (Applied Biosystems) results were exported as tab-delimited text files and imported into Microsoft Excel for further analysis.

Statistical data processing was performed using Excel and SPSS, and Pearson's correlation coefficients (r) were computed for each cross-platform comparison. Progression-free survival was measured from the day of surgery to the time of the first event of progression or to the last day of follow-up, according to the Kaplan-Meier method. Curves were compared using the log-rank test.

Results and Discussion

Selection of the reference gene

We chose the delta-delta Ct method [9] for real-time PCR measurement rather than using calibration curves. Although the delta-delta Ct method has stricter requirements, it can substantially reduce the number of PCR reactions.

The handling of gene expression data obtained by ATAC-PCR was similar to the handling of data from two-colour microarrays [12]. In both methods, the relative gene expression level compared to a control sample is measured and used for statistical analysis after data normalisation. In data normalisation of ATAC-PCR, each expression value was divided by the median of gene expression and then logarithmically converted. To choose the reference gene candidates whose expression was least changed between gliomas, we selected twelve genes exhibiting expression patterns that were most similar to the median gene expression pattern from 3,456 genes in the previous gene expression data matrix of 152 gliomas [5]. These twelve genes were expected to produce minimal variations in expression between

glioma samples. To select the best reference gene, the expression levels of the twelve genes were measured in 32 glioma samples using real-time PCR. The results are shown in Figure 1. Gene expression values are influenced by the amount of mRNA and the random variation caused by biological and experimental factors [13]. Because variation in the amount of mRNA was common to all of the genes, the difference in measurement was primarily due to the latter. The measurement of *UBL5* had the smallest variation; therefore, we selected it as the reference gene. Although the use of multiple reference genes is recommended by several reports [13,14], we chose a single reference gene for this case because the use of multiple reference genes would increase variations in measurement.

The first prognosis predictor developed for gliomas was based on the expression of 58 genes [5]. For the delta-delta Ct method, the amplification efficiency of a gene must be approximately equal to that of the reference gene. We performed real-time PCR amplification and fulfilled this criterion for 30 of the 58 genes. The original prognosis predictor classified gliomas into good and poor prognosis groups. The diagnostic scores calculated with the original 58 genes and the 30 genes chosen in this study had a high correlation (r = 0.95), and there was no difference between the classification results in the test set and those in the previous study [5]. Therefore, we decided to proceed with the 30 genes. A list of the genes and primer sequences is shown in Table 1.

Strategy for conversion

In our previous report of gene expression profiling of gliomas [5], we measured the relative expression levels against a control sample. Because the Ct value is inversely proportional to the amount of target nucleic acid present in the sample, the relative expression level of gene i of sample x, $er_i(x)$, is described as follows:

$$er_i(x) = (1+E)^{-(Ct_i(x)-Ct_i(c))}$$

Here, $Ct_i(x)$ and $Ct_i(c)$ are the Ct values of gene i of sample x and of the control sample, respectively. "$1+E$" represents the amplification efficiency of the real-time PCR, where $0 \le E \le 1$. The log-normalised gene expression, $en_i(x)$, is obtained by the following conversion:

$$\begin{aligned} en_i(x) &= \log(er_i(x) / er_{UBL5}(x)) \\ &= -\log(1+E) * (Ct_i(x) - Ct_{UBL5}(x)) \\ &\quad + \log(1+E) * (Ct_i(c) - Ct_{UBL5}(c)) \end{aligned}$$

Linear classifiers are most commonly used for molecular classification by gene expression profiles; an example is MammaPrint [2]. With a linear classifier, the diagnostic score is the sum of the normalised expression values multiplied by a coefficient determined from the learning data set. The diagnostic score of the prognosis predictor, the PC1 score, is described with Ct values as follows:

$$\begin{aligned} PC1(x) &= \sum_{i=1}^{n} a_i * en_i(x) \\ &= -\log(1+E) * \sum_{i=1}^{n} a_i * (Ct_i(x) - Ct_{UBL5}(x)) \\ &\quad + \log(1+E) * \sum_{i=1}^{n} a_i * (Ct_i(c) - Ct_{UBL5}(c)) \end{aligned}$$

Here, $PC1(x)$ is the PC1 score of sample x. "a_i" is a constant determined from the learning set in the previous study [5]. "n" is the number of diagnostic genes. $PC1(x)$ is alternatively described as follows, defining $PC1_{rt}(x)$ as the PC1 score of sample x measured by real-time PCR.

$$PC1(x) = \beta_1 * PC1_{rt}(x) + \beta_0$$

Here, $PC1_{rt}(x)$, β_1 and β_0 are as follows:

$$\begin{aligned} PC1_{rt}(x) &= \sum_{i=1}^{n} a_i * (Ct_i(x) - Ct_{UBL5}(x)) \\ \beta_1 &= -\log(1+E) \\ \beta_0 &= \log(1+E) * \sum_{i=1}^{n} a_i * (Ct_i(c) - Ct_{UBL5}(c)) \end{aligned}$$

Because the $PC1(x)$ value of the learning set was already determined, β_1 and β_0 can be determined by linear regression through measurement of $Ct_i(x)$ and

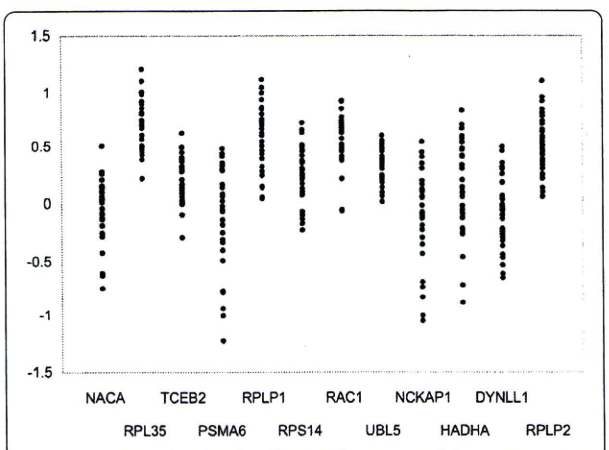


Figure 1 Expression levels of control gene candidates. Expression levels in 32 glioma tissues were measured and plotted for each gene.

Table 1 Primer sequences of the diagnostic genes

Gene Symbol	Forward	Reverse
IGFBP2	GCACATCCCCAACTGTGACA	TTCAGAGACATCTTGCACTGTTTG
VMP1	TGTCTTCTGTTGGGCTTGGAA	TGAGGCTATATGTGGACCCAGATA
MSN	GCCCCGGACTTCGTCTTC	AGGCCAAGATCCGCTTGTTA
TIMP1	CACAGACGGCCTTCTGCAAT	TGGTGTCCCCACGAACTTG
LGALS1	CTCCTGACGCTAAGAGCTTCGT	GAAGTGCAGGCACAGGTTGTT
CD63	CCCGAAAAACAACCACACTGC	GATGAGGAGGCTGAGGAGACC
NES	CAACAGCGACGGAGGTCTC	CCTCTACGCTCTCTTCTTTGAGT
CLIC1	TGTTCATGGTACTGTGGCTCAAG	GTCCGCCTTTTGGTGTCAAC
TNC	ACCACAATGGCAGATCCTTC	GCCTGCCTTCAAGATTTCTG
TAGLN2	CCTCTGGGAAGGAAAGAACATG	AGCCCACCCAGATTCATCAG
HES6	GACCAATGCCAGCCAGAG	GCAAGCCATCCATCAGAGG
VEGF	CCAAGGCCAGCACATAGGA	TCTTTGGTCTGCATTCACATTTG
VIM	TCCAAACTTTTCCTCCCTGAAC	GGGTATCAACCAGAGGGAGTGA
LDHA	CTGGGAGTTCACCCATTAAGCT	CAGGCACACTGGAATCTCCAT
RPIP8	CCCCCGTGGTCATCGA	GGTAGTCGTAGCTCTGCGTGAA
IFITM3	GGCTTCATAGCATTCGCCTACT	TCACGTCGCCAACCATCTT
PPIB	GGAGAGAAAGGATTTGGCTACAAA	CCTGGATCATGAAGTCCTTGATT
ALDOC	CGTCCGAACCATCCAGGAT	CCACACCCTTGTCAACCTTGAT
ZYX	CAGCAGCTAATGCAGGACATG	CAGAGTTCGTTGACAGCCACAT
UPAR	GTGTGTGGGTTAGACTTGTGCAA	AGGTAACGGCTTCGGGAATAG
LAMB2	CCACTGAAGGCGAGGTCATC	CCCGTAGGTTGGTGATCTTCAA
RTN1	CCGCATCTACAAGTCTGTTTTACAA	AAGCTCCAAGTAGGCCTTGAAAG
HMOX1	GGCAGAGAATGCTGAGTTCATG	AGGCCATCACCAGCTTGAAG
GM2A	GTCCCCCTGAGTTCTCCTCT	GCTCTTGGGCAGTGAGTAGG
S100A10	TGGAAAAGGAGTTCCCTGGAT	TACACTGGTCCAGGTCCTTCATT
BRSK2	GGAGGAGATGTCCAACCTGACA	AAGTTCCCAAACCAGGACTTCTT
MRCL3	AACAGAGATGGTTTCATCGACAAG	GTTGGATTCTTCCCCAATGAAG
GPX1	GCGGGGCAAGGTACTACTTA	CTCTTCGTTCTTGGCGTTCT
SOD2	AATCAGGATCCACTGCAAGGA	CGTGCTCCCACACATCAATC
RHOC	AATAAGAAGGACCTGAGGCAAGAC	ACGGGCTCCTGCTTCATCT
UBL5	AGCTGATTGCAGCCCAAACT	TCGTGTACCACTTCTTCAGGACAA

$Ct_{UBL5}(x)$ of the corresponding samples. The conversion formula would then be validated with the test set. It should be noted that this method does not require the use of a control sample (i.e., measurement of $Ct_i(c)$ and $Ct_{UBL5}(c)$).

Construction of the prognosis predictor based on real-time PCR

Using 44 samples from the learning set, we determined $PC1_{rt}(x)$ by measuring the Ct values. As expected, there was a high linear correlation between $PC1(x)$ and $PC1_{rt}(x)$ (r = 0.94), as shown in Figure 2.

We then measured the correlation in individual gene expression (Table 2) between the ATAC-PCR data (log-normalised) and the ΔCt values ($\Delta Ct(x) = Ct_i(x) - Ct_{UBL5}(x)$). The correlation for individual genes was less robust than that for the PC1 score: the correlation coefficients ranged from 0.6 to 0.9. These results suggest that the PC1 score could eliminate errors in measurement through the weighted averaging of gene expression. Because three genes (*VMP1, TNC and RHOC*) exhibited no correlation, we eliminated them from the diagnostic gene set. Because ATAC-PCR uses a single gene-specific primer designed for the 3' end of the mRNA, it may be less specific than conventional PCR using two primers. The absence of correlation may be due to the amplification of different genetic fragments or splicing variants. The parameters of the conversion formula were determined by linear regression (β_1, -0.37: β_0, -0.002).

Specific features of the expression of each gene may be obtained from the regression coefficient and intercept. Because the ATAC-PCR data were converted to a common logarithm during normalisation, the regression coefficient should be somewhere between zero and 0.30 (= $\log_{10}2$). In reality, the values ranged from 0.2 to 0.43, and ten genes demonstrated values exceeding 0.30. These results suggest a substantial degree of discrepancy between measurements obtained with ATAC-PCR and

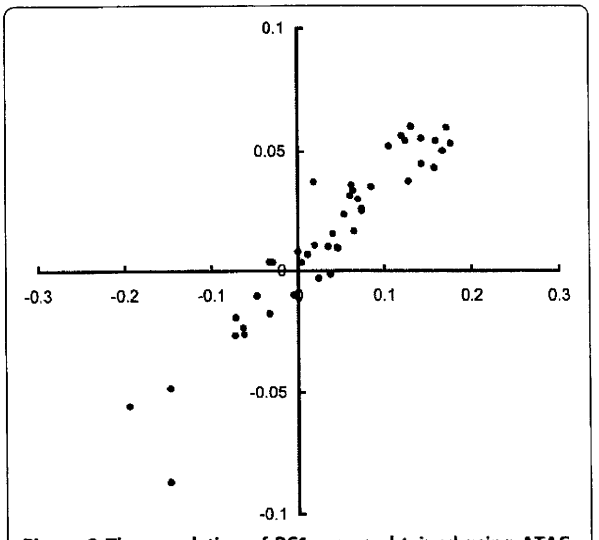


Figure 2 The correlation of PC1 scores obtained using ATAC-PCR and real-time PCR. Horizontal axis, PC1 score obtained with real-time PCR; vertical axis, PC1 score obtained with ATAC-PCR.

those determined using real-time PCR. The intercept indicates the general expression level of the gene; high intercept values indicate low levels of gene expression. With the exception of *VMP1*, the expression levels of the diagnostic genes were within two orders of magnitude of each other. The expression level of *UBL5* was in the middle range of all of the diagnostic genes.

Validation of the converted predictor

The converted predictor with 27 genes was validated with an additional sample set consisting of 16 samples from the previous test set [5] and 20 new samples. The samples were from anaplastic astrocytoma (grade III) or glioblastoma (grade IV). The PC1 score (*PC1(x)*) of each sample was calculated using ΔCt values measured using real-time PCR. The samples were classified into two prognosis groups with the threshold value set at zero, which was the threshold used in our previous study [5]. The performance of the classification was compared to conventional histopathological diagnosis. To have clinical utility, the predictor must have a classification ability superior to that of histopathological classification. The results of the Kaplan-Meier plot from the 36 samples revealed that the molecular classification was superior to histopathological diagnosis (log rank p-values, 0.023 and 0.137, respectively) (Figures 3A, B). The hazard ratio was 2.70 (95% confidence interval, 1.05-6.92) (p = 0.039) for molecular classification. No significant hazard ratio was obtained with histopathology (p = 0.16). We also noted that the classification results for the 16 samples from the original test set were the same as those previously obtained by ATAC-PCR. Thus, the new predictor based on real-time PCR is comparable to the previous predictor based on ATAC-PCR.

Table 2 Parameters for correlation between ATAC-PCR and real time PCR

gene name	correlation coefficient	regression coefficient	intercept
IGFBP2	0.90	0.27	0.32
VMP1	0.04	0.05	2.87
MSN	0.81	0.36	1.00
TIMP1	0.92	0.30	-0.31
LGALS1	0.85	0.36	-0.56
CD63	0.51	0.20	-0.52
NES	0.69	0.26	0.69
CLIC1	0.86	0.43	0.34
TNC	0.04	-0.02	-0.63
TAGLN2	0.66	0.34	0.13
HES6	0.77	0.29	0.60
VEGF	0.78	0.25	-0.11
VIM	0.77	0.30	-0.52
LDHA	0.73	0.33	-0.12
RPIP8	0.81	0.26	0.71
IFITM3	0.85	0.38	-0.75
PPIB	0.60	0.29	-0.10
ALDOC	0.73	0.28	-0.09
ZYX	0.68	0.36	0.54
UPAR	0.84	0.36	1.48
LAMB2	0.43	0.23	0.62
RTN1	0.82	0.29	0.66
HMOX1	0.87	0.30	0.62
GM2A	0.51	0.24	0.62
S100A10	0.79	0.28	-0.18
BRSK2	0.68	0.22	1.21
MRCL3	0.73	0.30	0.38
GPX1	0.70	0.33	-0.41
SOD2	0.74	0.31	0.23
RHOC	0.11	-0.08	-0.06

Further considerations

In the delta-delta Ct method, the selection of the reference gene is the most important technical point. It has been frequently noted that housekeeping genes are not necessarily adequate for use as reference genes [14,15] because of their variable expression levels. Although it is possible to use a combination of housekeeping genes [14], a reference gene or a set of reference genes selected from the expression data matrix of the target tissues is more desirable because the measurement of other tissues is not performed in diagnostic practice. We selected a reference gene from a set of genes exhibiting expression patterns that were similar to the median gene expression pattern for the glioma data. Alternative methods to select reference genes should also be applicable to the conversion method described here [13,16].

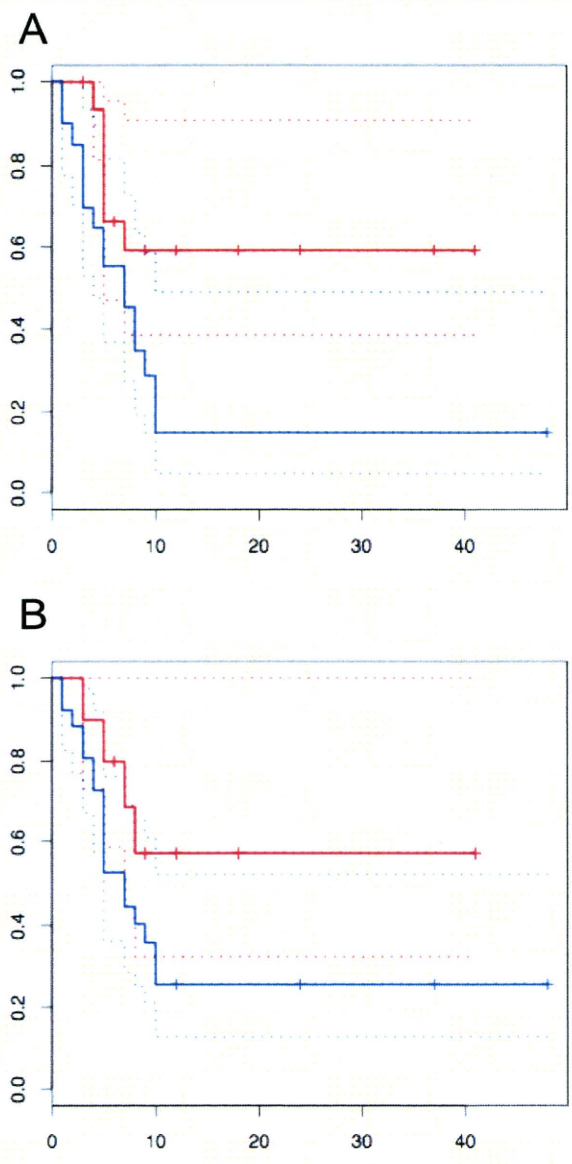


Figure 3 Kaplan-Meier analysis of grade III and grade IV glioma patients stratified either by (A) PC1 scores from real-time PCR data or by (B) histopathological diagnosis. Horizontal axis, month after diagnosis; vertical axis, progression-free survival probability. Blue lines, poor prognosis group (n = 20) (A) or grade IV (n = 26) (B); red lines, good prognosis group (n = 16) (A) or grade III (n = 10) (B). Log rank p-values were 0.023 (PC1 score) and 0.137 (histopathology). Dotted lines indicate 95% confidence intervals.

In the present study, the original classifier was developed from gene expression data obtained by ATAC-PCR. Our conversion method is based on the linear correlation between gene expression profiling data and ΔCt values. A linear correlation was observed between normalised microarray data and ΔCt values regardless of the normalisation procedure [17]. Thus, our method should also be applicable to linear classifiers obtained using microarrays. As described above, the correlation between diagnostic scores is higher than that between individual genes. As demonstrated by diagnostic tests for breast cancer, the scores calculated from multiple gene expression correlate with the biology (malignancy) much better than individual gene expression, which includes noise of biological and experimental origin. The higher correlation of diagnostic scores between the two PCR techniques is not surprising. This result suggests that the conversion should be performed with the diagnostic score; it is not appropriate to perform the conversion at the level of individual gene expression.

It should be noted that validation experiments were performed only for the conversion process and that the predictor itself is in the preliminary stages of development and still needs analytical clinical validation and clinical utility studies. In particular, because the original predictor may also be applicable for the prognosis prediction of grade II gliomas [5], the future cohort should include a large number of grade II gliomas. In grade II and III glioma patients, the optimal timing of radiation therapy is still controversial [18,19]. Precise risk assessment, including the ability to predict possible malignant transformation, may be useful for timing decisions and is the most promising feature of the new classification scheme.

Conclusions

We successfully converted a molecular classifier obtained by ATAC-PCR into a Ct value-based classifier. Our conversion procedure should also be applicable to linear classifiers developed from microarray data. Because errors in measurement are likely to be cancelled out during the calculation, the conversion of individual gene expression data is not an appropriate procedure. The predictor for gliomas is still in the preliminary stages of development and requires analytical clinical validation and clinical utility studies.

Acknowledgements
The authors thank Dr Shigeyuki Oba for advice on statistical analysis.

Author details
[1]Research Institute, Osaka Medical Center for Cancer and Cardiovascular Diseases, 1-3-3 Nakamichi, Higashinari-ku, Osaka, 537-8511, Japan. [2]Department of Neurosurgery, Kyoto University Graduate School of Medicine, 54 Kawahara-cho, Shogoin, Sakyo-ku, Kyoto-shi, Kyoto, 606-8507, Japan. [3]National Cerebral and Cardiovascular Center, 5-7-1 Fujishiro-dai, Suita, Osaka 565-8565, Japan. [4]Kitano Hospital. 2-4-20 Ohgimachi, Kita-ku, Osaka, 530-8480, Japan.

Authors' contributions
KK conceived and designed the study. SK performed the experimental work following advice from KT and YK. Statistical analysis was done by KK, MS and MK. MS, NM, NH, SM and JT recruited the glioma patients and were