REPRODUCIBILITY

Origin of the U87MG glioma cell line: Good news and bad news

Marie Allen,[1] Mia Bjerke,[1,2] Hanna Edlund,[1,3] Sven Nelander,[1] Bengt Westermark[1]*

Human tumor–derived cell lines are indispensable tools for basic and translational oncology. They have an infinite life span and are easy to handle and scalable, and results can be obtained with high reproducibility. However, a tumor-derived cell line may not be authentic to the tumor of origin. Two major questions emerge: Have the identity of the donor and the actual tumor origin of the cell line been accurately determined? To what extent does the cell line reflect the phenotype of the tumor type of origin? The importance of these questions is greatest in translational research. We have examined these questions using genetic profiling and transcriptome analysis in human glioma cell lines. We find that the DNA profile of the widely used glioma cell line U87MG is different from that of the original cells and that it is likely to be a bona fide human glioblastoma cell line of unknown origin.

INTRODUCTION

Many researchers rely on publicly available cell lines, such as those obtained from the American Type Culture Collection (ATCC) and other repositories. For research on human glioma, U87MG is probably the most commonly used cell line; searching with the keyword U87 glioma shows that there are currently more than 1700 entries in PubMed (more than 200 of which were published in 2015) and ~65,000 citations in the ISI Web of Science. The original U87MG cell line was established in our laboratory almost 50 years ago (*1*). Together with a few additional glioma cell lines, including the widely used U251MG and U373MG, the U87MG line was deposited in a cell repository managed by Fogh *et al.* (*2*) and later transferred to ATCC, from where the lines are currently available. Our long experience with the original U87MG (*3*) led us to question the authenticity of the ATCC cell line.

DNA fingerprinting, an important tool in forensic crime scene investigations and also used for authentication of cell lines, is often based on genotyping of short tandem repeat (STR) markers that are highly polymorphic. In a standard forensic DNA analysis, a multiplex polymerase chain reaction (PCR) amplification of up to 16 STRs are size-separated by capillary electrophoresis (*4*). Genotyping 16 STR markers results in a random match probability of around 1 in 2×10^{17} when using the PowerPlex 16 HS system (Promega) or the AmpFlSTR Identifiler PCR Amplification Kit (Applied Biosystems). A random match probability is the estimated frequency of a profile in a given population and reflects the probability that a randomly selected individual (or cell line) will have the exact same STR profile as the sample investigated.

RESULTS

The results from an STR profiling of U87MG (from both Uppsala and ATCC) are shown in Table 1. Twelve of 14 STR markers had different genotypes when we compared the data of U87MG from Uppsala to those of U87MG from ATCC. Thus, these two versions of U87MG do not originate from the same source. However, the data from the ATCC U87MG cell line were identical to previously published data by Bady *et al.* (*5*) and information from CLS Cell Lines Service (*5*, *6*). We then analyzed our entire panel of 19 additional human glioma cell lines that was established in our laboratory during the years 1966 to 1979. None of these cell lines was identical to the original U87MG cell line (table S1). The U118MG and U138 lines show identical profiles, as reported (*7*), indicating that cross-contamination has occurred. On the basis of the STR profiles presented in Table 1 and table S1, we concluded that the origin of the widely used U87MG line is different from that of the U87MG from Uppsala or any other glioma cell line established in our laboratory.

Table 1. Results for STR analysis of U87MG. Results from the PowerPlex 16 HS system are shown. STR genotypes are shown for the Uppsala versus ATCC versions of U87MG for 16 markers. The profile from previously published data of the U87MG cell line is also included (*5*). n.d., not determined.

DNA profile (STR)	U87MG Uppsala	U87MG ATCC	U87MG Bady *et al.*
CSF1PO	10/11	10/11	10/11
FGA	20/24	18/24	18/24
vWA	15/19	15/17	15/17
THO1	8/9	9.3/9.3	9.3/9.3
TPOX	8/11	8	8
Penta E	7/13	7/14	7/14
Penta D	11	9/14	9/14
D3S1358	15	16/17	16/17
D5S818	11/12	11/12	11/12
D7S820	10/13	8/9	8/9
D8S1179	13/14	10/11	10/11
D13S317	11/12	8/11	8/11
D16S539	9/13	12	12
D18S51	n.d.	n.d.	13
D21S11	28/30	28/32.2	28/32.2
Amel	X	X	X

[1]Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Uppsala University, SE-751 85 Uppsala, Sweden. [2]Department of Laboratory Medicine, Karolinska Institute, SE-141 86 Stockholm, Sweden. [3]Department of Organismal Biology, Uppsala University, SE-752 36 Uppsala, Sweden.
*Corresponding author. Email: bengt.westermark@igp.uu.se

Next, we attempted to confirm the glioma origin of the Uppsala U87MG cell line using mitochondrial DNA (mtDNA). Maternally inherited mtDNA is often used for forensic DNA profiling of damaged, degraded, and aged samples. The human mitochondrial genome is a circular double-stranded molecule of about 16,500 nucleotides in size and is located outside the nucleus. Each cell can contain up to 1000 copies of mtDNA. The high copy number per cell allows profiling even if only a small number of cells can be isolated. DNA was extracted from three sections of the original formalin-fixed, paraffin-embedded (FFPE) tumor tissue from 1966 (*8*), and the hypervariable (HV) regions HV1 and HV2 were amplified by PCR and analyzed by Sanger sequencing. A few nucleotides in the mtDNA from the FFPE tissue exhibited mixtures at some positions. Mixtures can reflect heteroplasmy or the sequence modifications that are typical of the damage that commonly occur when the samples are aged and/or formalin-fixed. The original tumor tissue and the Uppsala U87MG cell line have identical mtDNA profiles, whereas the ATCC and CLS variants of U87MG exhibit a different mtDNA profile (Table 2).

Together, our results show that the U87MG cell lines distributed by the ATCC and the CLS and the original U87MG cell line have different DNA profiles and that the mtDNA profile of the latter matches that of the original tumor tissue. The Uppsala version of U87MG is thus authentic. We do not know at which point the apparent mix-up occurred during the culturing of U87MG, but we sought to reveal the true tumor origin of the widely distributed commercial U87MG line.

Because the U87MG cell line we obtained from the ATCC and the CLS was not established from cells isolated from the same individual tumor as the Uppsala cell line, we explored its transcriptional similarity to a large database of cancer cell lines. For this analysis, we used the Affymetrix Gene 1.1 ST platform to obtain the transcript profile of the commercial U87MG line (*9*). We subsequently compared its profile to each of the cell lines in the Cancer Cell Line Encyclopedia (CCLE) of the Broad Institute of Massachusetts Institute of Technology and Harvard (*10*). The CCLE contains 1036 cell lines organized into 18 different tissue derivations [central nervous system (CNS), breast, and colon mucosa among others]. We counted the number of CCLE cases in which cells with a similar transcriptome are from a CNS tumor. In the commercial U87MG, five of the five most similar transcriptomes were indeed annotated as of CNS tumor origin. It is also instructive to sort all the CCLE cases on the basis of their transcriptional similarity to U87MG. Considering such a list (of length k, where k is an integer between 1 and 50) containing the CCLE cases most similar to U87MG, we computed the relative representation of different tissue origins. The analysis revealed a number of CNS origin cell lines that were higher than the random expectation ($P < 9.31\times 10^{-8}$ for $k > 5$, Fisher's exact test; Fig. 1). The analysis of the mRNA expression profile thus indicates that the commercially available

Table 2. Sequence variation in mtDNA from tumor tissue. The HV1 and HV2 regions of mtDNA from three different slides of FFPE sections from the original tumor were analyzed by Sanger sequencing. mtDNA profiles are also shown for the Uppsala, ATCC, and CLS versions of U87MG. The mtDNA variants were annotated with reference to the revised Cambridge Reference Sequence (rCRS). Experiments were performed after approval by the Regional Ethical Review Board (Dnr 2015/442).

mtDNA position	HV1	HV2				
	16,224	16,278	16,311	73	146	195
rCRS	**T**	**C**	**T**	**A**	**T**	**T**
U87MG FFPE section 1	T	C	T	A	T	T
U87MG FFPE section 2	T	C	T	A/G	T	T/C
U87MG FFPE section 3	T	C	T	A	T/C	T
U87MG Uppsala	T	C	T	A	T	T
U87MG ATCC	C	T	C	G	C	C
U87MG CLS	C	T	C	G	C	C

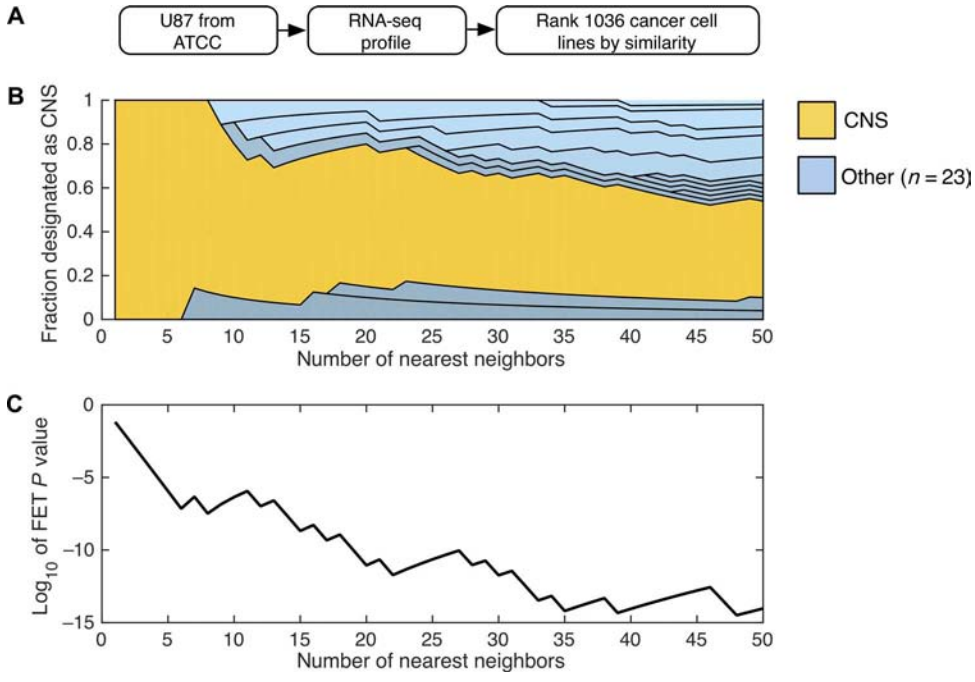

Fig. 1. U87MG from ATCC is of CNS origin. (**A**) Expression profile of U87MG cells from CLS obtained by RNA sequencing (RNA-seq), as compared to the CCLE compendium of 1036 cancer cell expression profiles (*10*). (**B**) Similarity analysis. The fraction of cell lines with a CNS origin among the k (x axis) most-similar lines. (**C**) Statistical assessment of tissue origin. Fisher's exact test was used to assess whether the number of CNS origin cell lines among the k (x axis) most-similar lines was above random expectation, given the total number of cell lines in CCLE, the total number of CNS-derived lines in CCLE, and the value of k. Log_{10} of the P value (y axis) shown for each k (negative values are more significant).

U87MG cell line is of CNS origin and is likely to be a bona fide human glioblastoma cell line, with an unknown patient origin.

DISCUSSION

The problem of the misidentification of cell lines has been highlighted in several reports (*11*, *12*), and a number of scientific journals now want the identity of the samples and cell lines used in their publication to be confirmed. Notably, *Nature* research journals, AACR publications, and some other scientific journals currently require cell identification by DNA analysis. This is an important step forward in quality assurance and cell line authentication. Nevertheless, it does not suffice to establish the identity of a cell line if the alleged origin of the reference line is unknown or even false, as we have shown in the case of U87MG. In general, a proper identification of a cell line requires a DNA profile that matches the tissue of origin, if circumstantial evidence for the origin of the cell line is not strong enough.

Cell line misidentification exacerbates another problem with established cell lines, that is, to which extent the cell lines are true representatives of the tumor of origin. This question has been addressed in the human glioma field. The seminal study of Lee *et al.* (*13*) showed that the transcriptomes of "classical" glioma cell lines grown in serum-containing medium, such as U87MG and U251MG, deviate from those of glioma tissues and are thus poor representatives of the tumor of origin. Cell lines established and maintained in serum-free neural stem cell medium retain tumor-specific phenotypes (*14*) and tumor-initiating capacity, and they are increasingly being used as experimental glioma models. Our laboratory has recently launched the Human Glioma Cell Culture resource as a publicly available repository of annotated and STR-genotyped glioma cell lines (*15*). Similar cell lines are also available from other laboratories. We strongly advocate the use of these or similar glioma cell lines in experimental and translational research on brain tumor biology. We suggest that the research community agree on the use of a set of such cell lines as references in research on cultured glioblastoma cells and that these verified cell lines replace U87MG and other classical cell lines.

MATERIALS AND METHODS

Study design

The overall objective of the study was to establish the origin of the widely used and publicly available U87MG glioblastoma cell line, using cells and tissue from the original U87MG tumor available in our laboratory. STR genotyping was used to identify cell lines, and analysis of HV regions of mtDNA was used to compare cell lines with FFPE tumor tissue. The STR typing was performed using two different commercial kits with most of the markers overlapping to account for technical variability. The PowerPlex 16 HS system was analyzed twice on different PCR products for biological replicates. Analysis of mtDNA was performed using two sets of primers and sequencing was performed from two separate PCRs for both HV regions. The mRNA expression profile of commercial U87MG cells was determined in a single experiment on a microarray platform for similarity analysis, followed by statistical assessment of tissue of origin.

Cell culture

Publicly available U87MG cultures were purchased from ATCC (LGC Standards GmbH) and CLS Cell Lines Service. All glioblastoma cell lines were established in our own laboratory during the years 1966 to 1979 and retrieved from frozen stocks. Cells were grown in Eagle's minimum essential medium supplemented with 10% fetal bovine serum and antibiotics.

STR typing of cell lines

A total of 15 STR loci and amelogenin for sex identification are included in the PowerPlex 16 HS system (Promega). The markers were amplified, according to the manufacturer's recommendations, by using 0.5 ng of input DNA. In addition, a total of 15 STRs and amelogenin included in the AmpFlSTR Identifiler PCR Amplification Kit (Applied Biosystems) were analyzed. The STR loci were amplified in a multiplex, according to the manufacturer's recommendations, using 1 ng of input DNA. Electrophoresis was performed using a 3730 genetic analyzer (Applied Biosystems). A mixture of 1.5 μl of the PCR product and 10 μl of Hi-Di formamide and GeneScan 500 LIZ Size Standard were loaded to the genetic analyzer. Allele calling was performed with the GeneMapper 4.0 software. The STR genotypes were interpreted independently by two different analysts.

mtDNA analysis of tumor tissue

Extraction of glioma cells from microscope slides from FFPE tissue samples was performed using the protocol described by Bus *et al.* (*8*). Analysis of mtDNA was performed by amplification of one 400–base pair (bp) fragment in the HV1 region (primers: F15971- TTAACTCCACCATTAGCACC and R16410-GAGGATGGTGGTCAAGGGAC) and one 373-bp fragment in the HV2 region (primers: F15-CACCCTATTAACCACTCACG and R429-CTGTTAAAAGTGCATACCGCCA) using the 9700 GeneAmp Thermocycler (Applied Biosystems). A total of 10 μl of DNA extract was added to 1× PCR Gold Buffer (Applied Biosytems), 2.4 mM $MgCl_2$, 0.2 μM of each primer, 5 U of AmpliTaq Gold DNA Polymerase, 0.2 mM of each deoxynucleotide triphosphate, 0.16 mg/ml bovine serum albumin, and 10% glycerol, in a total volume of 30 μl. The cycling conditions were 10 min at 95°C, followed by 38 cycles of 30 s at 94°C, 30 s at 60°C, and 30 s at 72°C, and a final extension step for 7 min at 72°C. Also, shorter mtDNA fragments were amplified with the HV1 primers 16128F and 16348R and the HV2 primers 45F and 287R (*16*). The shorter mtDNA products were amplified in a total volume of 25 μl with 10 μl of DNA extract and the KAPA2G Robust HotStart ReadyMix: 1× KAPA2G Robust HotStart ReadyMix (2 mM $MgCl_2$), with 0.4 μM of each primer. The cycling conditions were 3 min at 95°C, followed by 45 cycles of 15 s at 95°C, 15 s at 60°C, and 15 s at 72°C, and a final extension step for 7 min at 72°C. Purification of the PCR products was performed with 2 μl of ExoSap-IT and incubation at 37°C for 15 min and 80°C for 15 min for the inactivation of ExoSAP-IT. This was followed by Sanger sequencing using the BigDye Terminator version 3.1 Cycle Sequencing Kit, according to the manufacturer's protocol, and a 3730xl ABI Genetic Analyzer (Applied Biosystems). Sequencing was performed in both forward and reverse directions with the PCR primers as sequencing primers, and the sequence data were analyzed using the Sequencher 4.0.10 software (Gene Codes).

Computational analysis

We used an Affymetrix 1.1 ST platform to obtain the transcript profile of the U87MG line purchased from CLS (*7*). The CCLE data and tissue origin annotations were obtained from (*8*) as the distributed file "CCLE_expression_CN_muts_GENEE_2010-04-16.gctx." The pairwise transcriptional similarity between the U87MG transcript profile and the 1036 transcript profiles in CCLE was calculated using Pearson's

correlation coefficient r across distances of all 16,480 genes that were present in both data sets. The CCLE cell lines were then sorted in descending order on the basis of their r value. For a sequence of integers $k = 1,\ldots,50$, we divided the sorted list into two groups: (i) the top-ranking k cell lines and (ii) the bottom ranking (1036-k) cell lines. For each group [(i) and (ii)], we then counted the number of CCLE cell lines with a tissue origin annotation of CNS. We then applied Fisher's exact test to investigate whether we could reject the null hypothesis that the CNS annotation was randomly distributed between groups (i) and (ii). The resulting P values (one value for each k) are shown in Fig. 1.

SUPPLEMENTARY MATERIAL

www.sciencetranslationalmedicine.org/cgi/content/full/8/354/354re3/DC1
Table S1. STR analysis of the Uppsala collection of glioblastoma cell lines.

REFERENCES AND NOTES

1. J. Pontén, E. H. Macintyre, Long term culture of normal and neoplastic human glia. *Acta Pathol. Microbiol. Scand.* **74**, 465–486 (1968).
2. J. Fogh, W. C. Wright, J. D. Loveless, Absence of HeLa cell contamination in 169 cell lines derived from human tumors. *J. Natl. Cancer Inst.* **58**, 209–214 (1977).
3. B. Westermark, Growth regulatory interactions between stationary human glia-like cells and normal and neoplastic cells in culture. I. Normal cells. *Exp. Cell Res.* **81**, 195–206 (1973).
4. L. Roewer, DNA fingerprinting in forensics: Past, present, future. *Investig. Genet.* **4**, 22 (2013).
5. P. Bady, A.-C. Diserens, V. Castella, S. Kalt, K. Heinimann, M.-F. Hamou, M. Delorenzi, M. E. Hegi, DNA fingerprinting of glioma cell lines and considerations on similarity measurements. *Neuro Oncol.* **14**, 701–711 (2012).
6. CLS Cell Lines Service, http://clsgmbh.de.
7. American Type Culture Collection, http://lgcstandards-atcc.org/.
8. M. Bus, M. Allen, "DNA extraction from formalin fixed paraffin-embedded (FFPE) tissue samples," *The Molecular Methods Database*, www.molmeth.org [accessed 20 December 2013].
9. L. Schmidt, T. Kling, N. Monsefi, M. Olsson, C. Hansson, S. Baskaran, B. Lundgren, U. Martens, M. Häggblad, B. Westermark, K. F. Nilsson, L. Uhrbom, L. Karlsson-Lindahl, P. Gerlee, S. Nelander, Comparative drug pair screening across multiple glioblastoma cell lines reveals novel drug-drug interactions. *Neuro Oncol.* **15**, 1469–1478 (2013).
10. J. Barretina, G. Caponigro, N. Stransky, K. Venkatesan, A. A. Margolin, S. Kim, C. J. Wilson, J. Lehár, G. V. Kryukov, D. Sonkin, A. Reddy, M. Liu, L. Murray, M. F. Berger, J. E. Monahan, P. J. Morais, J. Meltzer, A. Korejwa, J. Jané-Valbuena, F. A. Mapa, J. Thibault, E. Bric-Furlong, P. Raman, A. Shipway, I. H. Engels, J. Cheng, G. K. Yu, J. Yu, P. Aspesi Jr., M. de Silva, K. Jagtap, M. D. Jones, L. Wang, C. Hatton, E. Palescandolo, S. Gupta, S. Mahan, C. Sougnez, R. C. Onofrio, T. Liefeld, L. MacConaill, W. Winckler, M. Reich, N. Li, J. P. Mesirov, S. B. Gabriel, G. Getz, K. Ardlie, V. Chan, V. E. Myer, B. L. Weber, J. Porter, M. Warmuth, P. Finan, J. L. Harris, M. Meyerson, T. R. Golub, M. P. Morrissey, W. R. Sellers, R. Schlegel, L. A. Garraway, The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012).
11. American Type Culture Collection Standards Development Organization Workgroup ASN-0002, Cell line misidentification: The beginning of the end. *Nat. Rev. Cancer* **10**, 441–448 (2010).
12. A. Capes-Davis, G. Theodosopoulos, I. Atkin, H. G. Drexler, A. Kohara, R. A. F. MacLeo, J. R. Masters, Y. Nakamura, Y. A. Reid, R. R. Reddel, R. I. Freshney, Check your cultures! A list of cross-contaminated or misidentified cell lines. *Int. J. Cancer* **127**, 1–8 (2010).
13. J. Lee, S. Kotliarova, Y. Kotliarov, A. Li, Q. Su, N. M. Donin, S. Pastorino, B. W. Purow, N. Christopher, W. Zhang, J. K. Park, H. A. Fine, Tumor stem cells derived from glioblastomas cultured in bFGF and EGF more closely mirror the phenotype and genotype of primary tumors than do serum-cultured cell lines. *Cancer Cell* **9**, 391–403 (2006).
14. S. M. Pollard, K. Yoshikawa, I. D. Clarke, D. Danovi, S. Stricker, R. Russell, J. Bayani, R. Head, M. Lee, M. Bernstein, J. A. Squire, A. Smith, P. Dirk, Glioma stem cell lines expanded in adherent culture have tumor-specific phenotypes and are suitable for chemical and genetic screens. *Cell Stem Cell* **4**, 568–580 (2009).
15. Y. Xie, T. Bergström, Y. Jiang, P. Johansson, V. D. Marinescu, N. Lindberg, A. Segerman, G. Wicher, M. Niklasson, S. Baskaran, S. Sreedharan, I. Everlien, M. Kastemar, A. Hermansson, L. Elfineh, S. Libard, E. C. Holland, G. Hesselager, I. Alafuzoff, B. Westermark, S. Nelander, K. Forsberg-Nilsson, L. Uhrbom, The human glioblastoma cell culture resource: Validated cell models representing all molecular subtypes. *EBioMedicine* **2**, 1351–1363 (2015).
16. M. Nilsson, G. Possnert, H. Edlund, B. Budowle, A. Kjellström, M. Allen, Analysis of the putative remains of a European patron saint–St. Birgitta. *PLOS One* **5**, e8986 (2010).

Acknowledgments: We thank M. Kastemar for technical help. **Funding:** This study was supported by the Swedish Cancer Society, the Swedish Crime Victim Compensation and Support Authority, and the Swedish strategic research program "eSSENCE." **Author contributions:** M.A., M.B., and H.E. performed DNA extraction from FFPE material, STR analysis, mtDNA analysis, and interpretation of data. S.N. performed the mRNA expression analysis and computational analysis. B.W. conceived the study and was responsible for the establishment and culture of human glioma cells. The manuscript was written by M.A., S.N., and B.W. All authors read and approved the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data not in the paper and the cell lines are available from the authors upon request.

Submitted 17 March 2016
Accepted 11 July 2016
Published 31 August 2016
10.1126/scitranslmed.aaf6853

Citation: M. Allen, M. Bjerke, H. Edlund, S. Nelander, B. Westermark, Origin of the U87MG glioma cell line: Good news and bad news. *Sci. Transl. Med.* **8**, 354re3 (2016).

Abstracts

One-sentence summary: The DNA profile of the publicly available and widely used glioma cell line U87MG is quite different from that of the original cell line.

Editor's Summary:
U87MG: Not what it used to be

A cell line commonly used for research on gliomas is found to be different from the original tumor from which it was derived. The authors, whose laboratory developed the cell line almost 50 years ago, compared the genetics of this line (obtained from ATCC) with those of the original tumor. They report that the DNA profile of the current cell line differs from that of the original cells but that it is likely to be a human glioblastoma cell line with unknown origins. This misidentification of a widely studied cell line reinforces the need for researchers to carefully validate the cell lines used in their research.