Discussion

This study shows molecular characteristics of lung cancers in never-smokers and smokers: (*i*) changes in expression of a relatively small number of miRNAs are involved in lung carcinogenesis in never-smokers; (*ii*) *EGFR* mutations may reinforce some of these changes in miRNA expression, e.g., an increase in miR-21; (*iii*) miR-138 on 3p21.33, a chromosomal region carrying a long-sought lung cancer suppressor gene, is down-regulated preferentially in never-smoker cases; and (*iv*) miR-21 is one of the most aberrantly increased miRNAs in both never-smoker and smoker cases. There was no significant difference in the expression levels of miR-21 when stage I cases (n = 21) were compared with stage II, III, and IV cases (n = 7) (data not shown), suggesting that increased miR-21 expression is an early event in lung carcinogenesis. These findings identified miR-21 as a major miRNA that may play an oncogenic role in lung carcinogenesis and prompted us to choose it as a candidate for molecular targets in treatment of lung cancers in never-smokers as well as in smokers. Given the relationship between *EGFR* mutations and miR-21 up-regulation, we also hypothesized that this miRNA might have implications in improving EGFR-TKI therapy, whose effectiveness is correlated with *EGFR* gene status and smoking history of the patients (2, 4–6, 31).

Although high levels of miR-21 expression have been reported in various types of human tumors (19, 30, 32), including lung cancer from both smokers (18) and never-smokers (this study), the mechanism that up-regulates miR-21 during carcinogenesis is not well understood. In addition to the miRNA microarray data showing higher levels of miR-21 in *EGFR*-mutant cases (Table 3), the in vitro analyses using NSCLC cell lines showed that the activated EGFR signaling up-regulates miR-21 expression. A statistically significant positive correlation was observed between miR-21 expression levels and p-EGFR levels in NSCLC cell lines (Fig. 1*B*). Furthermore, the treatment with the EGFR-TKI (AG1478) inhibited miR-21 expression in 2 NSCLC cell lines with elevated p-EGFR, *EGFR*-mutant H3255 (Fig. 2) and *EGFR* wild-type H441 (Fig. S4), providing a mechanistic link between the activated EGFR signaling pathway and the aberrant up-regulation of miR-21 and a therapeutic basis for inhibition of miR-21 in lung cancers with EGFR activation. STAT3 reportedly signals IL-6–induced up-regulation of miR-21 in multiple myeloma cells (33). However, siRNA-mediated knockdown of endogenous STAT3 expression did not affect miR-21 levels in H441, H1650, and H1975 cells (data not shown), suggesting that STAT3 does not play a primary role in the EGFR signaling-induced up-regulation of miR-21 in NSCLC cells. It remains to be examined whether activator protein-1 (AP-1), which is activated by the EGFR signaling (34) and activates the miR-21 transcription through the binding to the promoter (35), is responsible for the increased expression of miR-21 in NSCLC cells. Nevertheless, there also should be EGFR-independent mechanisms to control miR-21 expression, because miR-21 was expressed abundantly in A549 cells without *EGFR* mutation or p-EGFR (Figs. 1*A* and S3*B*), and no increased miR-21 expression was observed in H1650 cells with *EGFR* mutation and increased p-EGFR (Figs. 1*A* and S3*C*). An increased copy number or amplification of the chromosomal region carrying *miR-21* (17q23.1) (36) may be an EGFR-independent mechanism.

Antisense oligonucleotide-mediated knockdown of miR-21 induced or enhanced apoptotic responses in 2 NSCLC cell lines, H3255 and H441 (Fig. 3) probably recapitulating some lung cancer cases from never-smokers. H3255 and H441 both expressed elevated levels of miR-21 (Fig. 1*A*) but had biologically and genetically different features. H3255 was highly responsive to EGFR-TKI (Fig. 1*A*), expressed high levels of p-EGFR (Figs. 2*A* and S3*C*), and had mutated and amplified *EGFR* (Fig. 1*A*) (37). H441 was less responsive to EGFR-TKI (Fig. 1*A*), expressed low levels of p-EGFR (Fig. S3*B*), and had wild-type *EGFR* (Fig. 1*A*). As previously reported in other types of human cancer cells (32, 33), the antisense inhibition of miR-21 by itself led to increased apoptotic response in H441 cells (Fig. 3 *C* and *D*), suggesting that miR-21 also can be a therapeutic target in lung cancers. Importantly, in both cell lines, anti-miR-21 significantly enhanced the apoptotic response induced by AG1478 (Fig. 3 *B* and *C*). The lack of effect of anti-miR-21 alone in H3255 cells (Fig. 3*B*) may indicate that the combinational use of anti-miR-21 and EGFR-TKI is required to attenuate effectively the constitutively activated EGFR signaling pathway to cell survival, which is evidenced by the highest levels of p-EGFR (Fig. S3*C*) and miR-21 (Fig. 1*A*). Although EGFR-TKIs are in wide clinical use for lung cancer (38), and inhibition of oncogenic miRNAs is a new promising approach in cancer therapy (39), this study reveals that the combination of these 2 therapeutic strategies can be significantly more effective than either alone. The finding is of particular importance in developing better treatment for lung cancer patients of non-Asian ethnicity, who tend to be less responsive to the current EGFR-TKI therapy (40). This study also has potential clinical implications in preventing and rescuing acquired EGFR-TKI resistance in NSCLC, an issue with important clinical relevance. Besides a secondary T790M mutation (41) and acquired MET amplification (42), selection of an *EGFR* wild-type subpopulation on a background of wild-type/mutant mixture leads to acquired EGFR-TKI resistance in NSCLC (43). The combinatorial use of EGFR-TKI and anti-miR-21 could prevent and rescue such acquired resistance caused by the selection for wild-type *EGFR*, because anti-miR-21 is effective on both *EGFR* wild-type and mutant tumor cells. Successful i.v. administration of a locked nucleic acid-modified antisense miRNA in primates (44) supports the feasibility of in vivo targeting miRNAs in therapy of human diseases.

Last, our lists of dysregulated miRNAs (Tables 2 and 3) include a number of other miRNAs that may have an oncogenic or tumor-suppressive role in lung carcinogenesis, e.g., up-regulated miR-141 (45), up-regulated miR-210 (46), down-regulated miR-126 (47), and down-regulated miR-486 (48). Further studies will address the roles of other individual miRNAs in lung carcinogenesis and a possible therapeutic relevance of overexpression and/or knockdown of multiple miRNAs.

In conclusion, lung cancers in never-smokers have a characteristic profile of miRNA expression. MiR-21 is a downstream effector of the activated EGFR signaling pathway and can be a therapeutic target in lung cancers with and without *EGFR* mutations. Antisense inhibition of miR-21 may improve clinical response to EGFR-TKI therapy.

Materials and Methods

Clinical Samples and Cell Lines. Lung cancer tissues and corresponding noncancerous tissues were obtained from never-smokers who had undergone surgical resection between 2000 and 2004 at the University of Maryland Medical Center (n = 15) or the Mayo Clinic (n = 7) in United States or at the Hamamatsu University School of Medicine (n = 6) in Japan (Tables 1 and S1). Institutional review board approval and written informed consent from all patients were obtained at each collection site. H3255 was provided by Dr. Bahar of the National Cancer Institute. A549, H23, H441, H1650, H1975, H157, and H1299 were purchased from American Type Culture Collection (ATCC). *hTERT*-immortalized normal human bronchial epithelial cells (HBET2) were established in our laboratory.

Microarray Analysis. Microarray analysis was performed as previously described (21, 49). Briefly, 5 μg of total RNA was used for hybridization on miRNA microarray chips containing 389 probes in triplicate (Ohio State microRNA microarray version 3.0). Processed slides were scanned using a PerkinElmer ScanArray XL5K Scanner. With statistical software R (http://www.r-project.org/), only spot values that were not flagged by the image quantification software GenePix Pro 6.0.1.00 (www.moleculardevices.com/pages/software/gn_genepix_pro.html) and whose foreground intensities were larger than background intensities were used. The remaining spots then were normalized by locally weighted scatterplot smoothing (LOESS), and duplicate spots were averaged. The preprocessed and normalized data then were imported into BRB-ArrayTools version 3.5.0 (http://linus.nci.nih.gov/BRB-ArrayTools.html). Finally, 291 miRNAs with non-missing log values present in more than 75% of the samples were selected. More information is available in *SI Materials and Methods*.

Real-Time RT-PCR Analysis. Expression of mature miRNAs was examined by qRT-PCR analysis using a TaqMan Human MicroRNA Assay kit and a PRISM 7700 Sequence Detector System (Applied Biosystems). RNU6B was an endogenous control (#4373381, Applied Biosystems). Gene expression data (mean ± SD from triplicate samples) were shown as $2^{-\Delta\Delta CT}$ (50).

Cell Treatment and Growth Inhibition Assay. To evaluate the effect of AG1478 on the EGFR signaling pathway and miR-21 expression levels, lung cancer cell lines were serum starved for 24 h, incubated in the presence or absence of AG1478 (2 μM or 10 μM; Calbiochem) for 2 h, and then for an additional 15 min in the presence or absence of EGF (20 ng/mL; Promega). Growth inhibition was assessed by 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl) 2H-tetrazolium, inner salt (MTS) assay (Dojindo). Cells (5,000/well) were seeded into 96-well plates, and increasing concentrations of AG1478 (0, 0.4, 2.0, 10, and 50 μM) were added. After incubation for 72 h at 37 °C, MTS was added and incubated for 2 h at 37 °C; then absorbance at 450 nm was measured. The IC_{50} value was defined as the concentration needed for 50% reduction of the growth.

Western Blot Analysis. Cells were lysed in buffer containing 50 mM Tris-HCl, pH 7.6, 150 mM NaCl, 0.1% SDS, 1% Nonidet P-40, and 0.5% sodium-deoxycholate. Ten μg of proteins were separated by gel electrophoresis on 10% gels, transferred to nitrocellulose membranes, and detected by immunoblotting using a chemiluminescence system (GE Healthcare Bio-Sciences Corp.). The images were quantified using National Institutes of Health ImageJ1.40g (http://rsb.info.nih.gov/ij/). Antibodies are in *SI Materials and Methods*.

Oligonucleotide Transfection and Apoptosis Assay. Cells were transfected with the oligonucleotides (at a final concentration of 40 nM) using LipofectAMINE 2000 reagent (Invitrogen). After 72 h, the cells were incubated in the presence or absence of 0.2 μM of AG1478 for 24 h (H3255) or 10 μM of AG1478 for 72 h (H441). Activities of caspase-3 and caspase-7 were analyzed using ApoONE Homogeneous Caspase 3/7 Assay (Promega). Each experiment was done in triplicate and at least 4 times independently. The data were shown as mean ± SD. Oligonucleotide sequences are in *SI Materials and Methods*.

Statistical Analysis. The paired *t*-test identified miRNAs that were differentially expressed in lung cancer tissues and normal lung tissues ($P < 0.01$, FDR < 0.15). We also identified miRNAs that were differently expressed in *EGFR*-mutant and wild-type lung cancers using the F-test ($P < 0.01$, FDR < 0.15). The paired *t*-test was used to analyze differences in miRNA expression in tumors and corresponding normal tissues for qRT-PCR data. Graphpad Prism v5.0 (Graphpad Software Inc.) analysis was used for the Pearson's correlation. All statistical tests were 2-sided, and statistical significance was defined as $P < 0.05$.

ACKNOWLEDGMENTS. We thank Dr. Kensuke Kumamoto for helpful discussions, Drs. Raymond T. Jones, Andrew Borkowski, and Mark J. Krasna for sample collection and pathology reports, and Audrey Salabes for interviewing the lung cancer patients. This work was supported by the Intramural Research Program of the National Institutes of Health, National Cancer Institute, and Center for Cancer Research.

1. Parkin DM, Bray F, Ferlay J, Pisani P (2005) Global cancer statistics, 2002. *CA Cancer J Clin* 55:74–108.
2. Mountzios G, Fouret P, Soria JC (2008) Mechanisms of disease: Signal transduction in lung carcinogenesis—a comparison of smokers and never-smokers. *Nature Clinical Practice Oncology* 5:610–618.
3. Hollstein M, Sidransky D, Vogelstein B, Harris CC (1991) p53 mutations in human cancers. *Science* 253:49–53.
4. Vähäkangas KH, et al. (2001) p53 and *K-ras* mutations in lung cancers from former and never-smoking women. *Cancer Res* 61:4350–4356.
5. Mounawar M, et al. (2007) Patterns of *EGFR*, *HER2*, *TP53*, and *KRAS* mutations of $p14^{arf}$ expression in non-small cell lung cancers in relation to smoking history. *Cancer Res* 67:5667–5672.
6. Sun S, Schiller JH, Gazdar AF (2007) Lung cancer in never smokers—a different disease. *Nature Reviews Cancer* 7:778–790.
7. Lynch TJ, et al. (2004) Activating mutations in the epidermal growth factor receptor. *N Engl J Med* 350:2129–2139.
8. Paez JG, et al. (2004) EGFR mutations in lung cancer: Correlation with clinical response to gefitinib therapy. *Science* 304:1497–1500.
9. Uramoto H, Mitsudomi T (2007) Which biomarker predicts benefit from EGFR-TKI treatment for patients with lung cancer? *Br J Cancer* 96:857–863.
10. Calin GA, et al. (2002) Frequent deletions and down-regulation of micro-RNA genes *miR15* and *miR16* at 13q14 in chronic lymphocytic leukemia. *Proc Natl Acad Sci USA* 99:15524–15529.
11. Inamura K, et al. (2007) let-7 microRNA expression is reduced in bronchioloalveolar carcinoma, a non-invasive carcinoma, and is not correlated with prognosis. *Lung Cancer* 58:392–396.
12. Lu J, et al. (2005) MicroRNA expression profiles classify human cancers. *Nature* 435:834–838.
13. Calin GA, et al. (2005) A microRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *N Engl J Med* 353:1793–1801.
14. Volinia S, et al. (2006) A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc Natl Acad Sci USA* 103:2257–2261.
15. Johnson SM, et al. (2005) RAS is regulated by the let-7 microRNA family. *Cell* 120:635–647.
16. Takamizawa J, et al. (2004) Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. *Cancer Res* 64:3753–3756.
17. Hayashita Y, et al. (2005) A polycistronic microRNA cluster, miR-17–92, is overexpressed in human lung cancers and enhances cell proliferation. *Cancer Res* 65:9628–9632.
18. Yanaihara N, et al. (2006) Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. *Cancer Cell* 9:189–198.
19. Schetter AJ, et al. (2008) MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma. *J Am Med Assoc* 299:425–436.
20. Weidhaas JB, et al. (2007) MicroRNAs as potential agents to alter resistance to cytotoxic anticancer therapy. *Cancer Res* 67:11111–11116.
21. Liu CG, Calin GA, Volinia S, Croce CM (2008) MicroRNA expression profiling using microarrays. *Nature Protocols* 3:563–578.
22. Landi MT, et al. (2008) Gene expression signature of cigarette smoking and its role in lung adenocarcinoma development and survival. *PLoS ONE* 20:e1651.
23. Ding L, et al. (2008) Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* 455:1069–1075.
24. Zabarovsky ER, Lerman MI, Minna JD (2002) Tumor suppressor genes on chromosome 3p involved in the pathogenesis of lung and other cancers. *Oncogene* 21:6915–6935.
25. Mitomo S, et al. (2008) Downregulation of miR-138 is associated with overexpression of human telomerase reverse transcriptase protein in human anaplastic thyroid carcinoma cell lines. *Cancer Sci* 99:280–286.
26. Horikawa I, Barrett JC (2003) Transcriptional regulation of the telomerase *hTERT* gene as a target for cellular and viral oncogenic mechanisms. *Carcinogenesis* 24:1167–1176.
27. Sharma SV, Bell DW, Settleman J, Haber DA (2007) Epidermal growth factor receptor mutations in lung cancer. *Nat Rev Cancer* 7:169–181.
28. Tracy S, et al. (2004) Gefitinib induces apoptosis in the $EGFR^{L858R}$ non-small-cell lung cancer cell line H3255. *Cancer Res* 64:7241–7244.
29. Sordella R, Bell DW, Haber DA, Settleman J (2004) Gefitinib-sensitizing EGFR mutations in lung cancer activate anti-apoptotic pathways. *Science* 305:1163–1167.
30. Meng F, et al. (2007) MicroRNA-21 regulates expression of the PTEN tumor suppressor gene in human hepatocellular cancer. *Gastroenterology* 133:647–658.
31. Toyooka S, et al. (2007) The impact of sex and smoking status on the mutational spectrum of epidermal growth factor receptor gene in non small cell lung cancer. *Clin Cancer Res* 13:5763–5768.
32. Chan JA, Krichevsky AM, Kosik KS (2005) MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells. *Cancer Res* 65:6029–6033.
33. Löffler D, et al. (2007) Interleukin-6 dependent survival of multiple myeloma cells involves the Stat3-mediated induction of microRNA-21 through a highly conserved enhancer. *Blood* 110:1330–1333.
34. Li J, et al. (2003) Differential requirement of EGF receptor and its tyrosine kinase for AP-1 transactivation induced by EGF and TPA. *Oncogene* 22:211–219.
35. Fujita S, et al. (2008) miR-21 Gene expression triggered by AP-1 is sustained through a double-negative feedback mechanism. *J Mol Biol* 378:492–504.
36. Łastowska M, et al. (2002) Breakpoint position on 17q identifies the most aggressive neuroblastoma tumors. *Genes Chromosomes Cancer* 34:428–436.
37. Engelman JA, et al. (2006) Allelic dilution obscures detection of a biologically significant resistance mutation in EGFR-amplified lung cancer. *J Clin Invest* 116:2695–2706.
38. Shepherd FA, et al. (2005) Erlotinib in previously treated non-small-cell lung cancer. *N Engl J Med* 353:123–132.
39. Bumcrot D, Manoharan M, Koteliansky V, Sah DW (2006) RNAi therapeutics: A potential new class of pharmaceutical drugs. *Nat Chem Biol* 2:711–719.
40. Thatcher N, et al. (2005) Gefitinib plus best supportive care in previously treated patients with refractory advanced non-small-cell lung cancer: Results from a randomised, placebo-controlled, multicentre study (Iressa Survival Evaluation in Lung Cancer). *Lancet* 366:1527–1537.
41. Kobayashi S, et al. (2005) EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *N Engl J Med* 24:786–792.
42. Engelman JA, et al. (2007) MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science* 316:1039–1043.
43. Jiang SX, et al. (2008) EGFR genetic heterogeneity of nonsmall cell lung cancers contributing to acquired gefitinib resistance. *Int J Cancer* 123:2480–2486.
44. Elmén J, et al. (2008) LNA-mediated microRNA silencing in non-human primates. *Nature* 452:896–899.
45. Iorio MV, et al. (2007) MicroRNA signatures in human ovarian cancer. *Cancer Res* 67:8699–8707.
46. Foekens JA, et al. (2008) Four miRNAs associated with aggressiveness of lymph node-negative, estrogen receptor-positive human breast cancer. *Proc Natl Acad Sci USA* 105:13021–13026.
47. Liu B, et al. (2009) MiR-126 restoration down-regulate VEGF and inhibit the growth of lung cancer cell lines in vitro and in vivo. *Lung Cancer*, in press.
48. Gal H, et al. (2008) MIR-451 and imatinib mesylate inhibit tumor growth of glioblastoma stem cells. *Biochem Biophys Res Commun* 376:86–90.
49. Liu CG, et al. (2004) An oligonucleotide microchip for genome-wide microRNA profiling in human and mouse tissues. *Proc Natl Acad Sci USA* 101:9740–9744.
50. Schmittgen TD, et al. (2008) Real-time PCR quantification of precursor and mature microRNA. *Methods* 44:31–38.

Int. J. Cancer: **125,** 973–976 (2009)

SHORT REPORT

Identification and characterization of a novel germline *p53* mutation in a patient with glioblastoma and colon cancer

Hidetaka Yamada[1], Kazuya Shinmura[1], Yasuhiro Yamamura[2], Kiyotaka Kurachi[3], Toshio Nakamura[3], Toshihiro Tsuneyoshi[4], Naoki Yokota[2], Masato Maekawa[5] and Haruhiko Sugimura[1]*

[1]*First Department of Pathology, Hamamatsu University School of Medicine, Hamamatsu, Japan*
[2]*Department of Neurosurgery, Hamamatsu University School of Medicine, Hamamatsu, Japan*
[3]*Second Department of Surgery, Hamamatsu University School of Medicine, Hamamatsu, Japan*
[4]*Department of Materials and Life Science, Shizuoka Institute of Science and Technology, Fukuroi, Japan*
[5]*Department of Laboratory Medicine, Hamamatsu University School of Medicine, Hamamatsu, Japan*

Germline mutations in the *p53* tumor suppressor gene have been identified in patients with Li-Fraumeni syndrome (LFS) and patients with Li-Fraumeni-like syndrome (LFL). However, to date, germline *p53* mutations in patients not fulfilling the criteria of LFS or LFL have been reported only very rarely. In our study, a novel germline c.584T>C (p.Ile195Thr) mutation of the *p53* gene was found in a 21–year-old male with a glioblastoma and colon cancer. He had no family history of cancer within second-degree relatives, and loss of the wild-type *p53* allele and overexpression of p53 protein were observed in both tumors. Functional analyses revealed transactivation and growth suppressive function activities of the Thr195-type p53 to be impaired. These results suggest germline *p53* mutations to possibly be responsible for a subset of young adult patient with multiple malignant tumors, even those not meeting the clinical criteria for LFS or LFL.

Key words: p53; germline mutation; glioblastoma; colon cancer; adolescent and young adult (AYA)

Li-Fraumeni syndrome (LFS) is an inherited cancer predisposition syndrome characterized by a wide spectrum of neoplasms occurring at a young age.[1] Germline mutations in the *p53* tumor suppressor gene have been identified in patients with LFS and patients with Li-Fraumeni-like syndrome (LFL).[2–5] However, to date, germline *p53* mutations in patients not fulfilling the criteria of LFS or LFL have been reported only very rarely.[6] In the present study, we identified a novel germline *p53* mutation in a patient with a glioblastoma and colon cancer, whose family history was not consistent with either LFS or LFL, examined *p53* allelic status and p53 protein expression in the tumors and evaluated the functional activities of the novel mutant p53.

Material and methods

Samples, cell lines and ethical considerations

Samples of the patient's colon, brain tissue and blood were obtained at surgery at Hamamatsu University Hospital. DNA was extracted from all samples. The human H1299 cell line was purchased from the American Type Culture Collection (Manassas, VA). The research protocol was approved by the Institutional Review Board of Hamamatsu University School of Medicine.

Polymerase chain reaction (PCR) and sequencing analysis

Fragments covering all coding exons and boundary regions of the *p53* gene were amplified by PCR and sequenced as described previously.[7,8]

Genotyping of the c.584T>C mutation

The c.584T>C mutation was genotyped by PCR-restriction fragment length polymorphism (RFLP) analysis using the *Ava*I enzyme.

Loss of heterozygosity analysis

Fluorescence-based PCR was used to examine the loss of heterozygosity (LOH) at the *p53* locus as described previously.[9]

Immunohistochemical analysis

Tissue slices were stained with hematoxylin-eosin or immunostained with DO-7 (DAKO, Kyoto, Japan), an anti-p53 monoclonal antibody.[10]

Luciferase reporter assay

p53-null human H1299 cells were transiently cotransfected with the p53 expression vector, firefly luciferase reporter vector pGL4.10 (Promega, Madison, WI) containing a fragment of the *p21* (*CDKN1A*), *BAX* or *MDM2* promoter and a Renilla luciferase reporter vector pGL4.74 (Promega), and 24 hr after transfection, luciferase activity was measured by a Dual-Luciferase Reporter Assay System (Promega) as described previously.[7]

Colony formation assay

p53-null H1299 cell lines were transfected with the p53 expression vector using Lipofectamine 2000 reagent (Invitrogen, Carlsbad, CA), followed by selection in media containing G418 antibiotics at 1 mg/ml for 2 weeks. Colonies were fixed, stained and counted.

Statistical analysis

Statistical comparisons were performed by the two-tailed Student's *t* test with Excel software (Microsoft Corp., Redmond, WA).

Results and discussion

A male patient was diagnosed with a glioblastoma at 20 years of age and died of colon cancer at 21 years of age (Fig. 1*a*). First, we considered the possibility of an inherited abnormality of mismatch repair genes and performed a microsatellite instability analysis of tissue from both tumors. However, both tumors were found to be microsatellite-stable (data not shown), suggesting no involvement of mismatch repair gene abnormalities. Next,

Grant sponsor: Ministry of Health, Labour and Welfare; Grant number: 19-19; Grant sponsor: Japan Society for the Promotion of Science; Grant number: 19790286; Grant sponsor: Ministry of Education, Culture, Sports, Science and Technology; Grant number: 18014009; Grant sponsor: 21st Century Center of Excellence Program.
***Correspondence to:** First Department of Pathology, Hamamatsu University School of Medicine, 1-20-1 Handayama, Higashi Ward, Hamamatsu, Shizuoka 431-3192, Japan. Fax: +81-53-435-2225.
E-mail: hsugimur@hama-med.ac.jp
Received 13 November 2008; Accepted after revision 27 February 2009
DOI 10.1002/ijc.24432
Published online 19 March 2009 in Wiley InterScience (www.interscience.wiley.com).

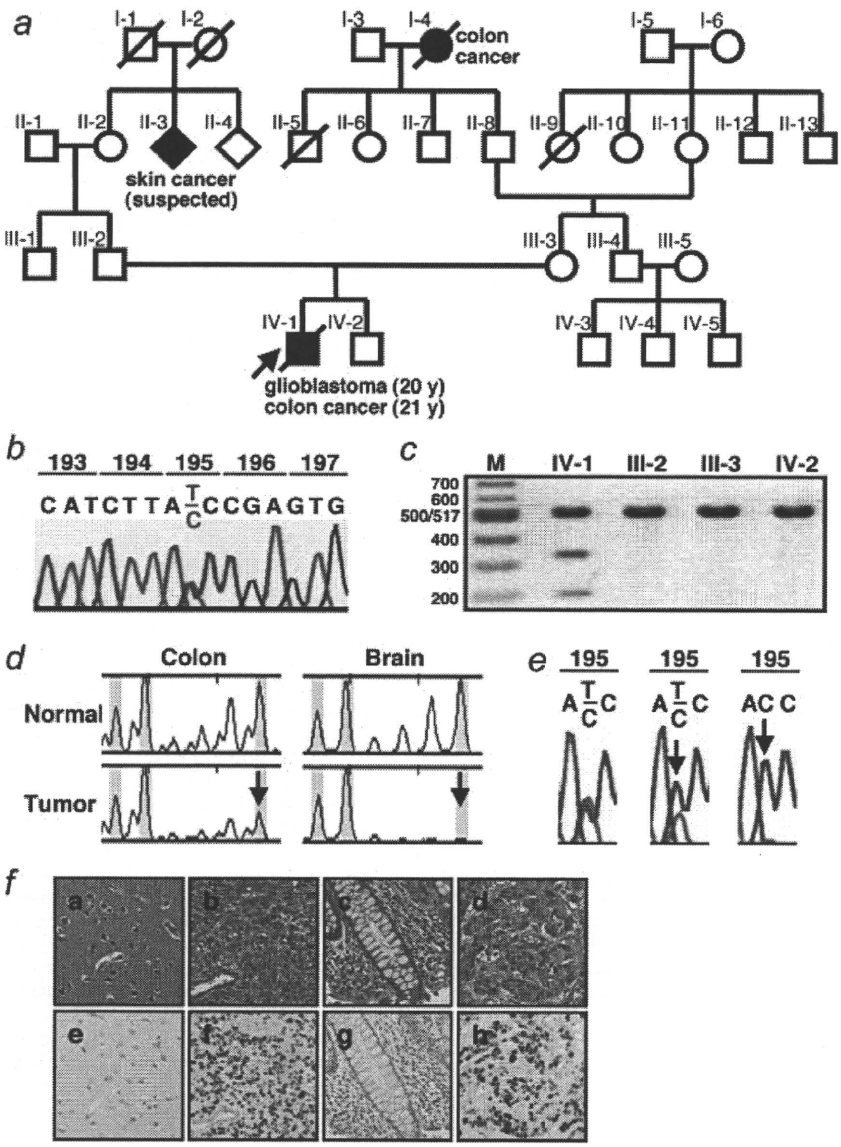

FIGURE 1 – Identification of the c.584T>C (p.Ile195Thr) germline mutation of the *p53* gene in a patient with glioblastoma and colon cancer. (*a*) Pedigree of the patient's family. Square, male; circle, female. Solid symbols represent cancer patients. Symbols with a slash indicate deceased individuals. The arrow points to the proband. (*b*) Detection of a germline *p53* mutation in blood DNA from this case. Direct sequencing analysis of exon 6 of the *p53* gene revealed a heterozygous mutation with an ATC allele and ACC allele at codon 195. (*c*) Genotyping of the p.Ile195Thr mutation in blood DNA from patient and his relatives by PCR-RFLP analysis. Individuals III-2, III-3, IV-1 and IV-2 are shown in (*a*). M, DNA size marker. The size of the PCR product was 527 bp. A mutant allele was digested into 2 fragments of 333 bp and 194 bp with an *Ava*I restriction enzyme, but the wt allele was not. (*d*) LOH analysis at the *p53* locus using a TP53 microsatellite marker. Fluorescent PCR products were analyzed using an ABI PRISM 310 Genetic Analyzer (Applied Biosystems). The arrow indicates loss of 1 allele. (*e*) LOH analysis at the *p53* gene locus by sequencing of the mutation site at codon 195. Electropherograms of normal colonic mucosa, colon cancer and glioblastoma are shown at the left, middle and right, respectively. The arrow indicates loss of the wt allele. (*f*) Immunohistochemical analyses of p53 protein in the glioblastoma and colon cancer. (*a*, *e*), normal brain; (*b*, *f*) glioblastoma; (*c*, *g*) normal colon; (*d*, *h*) colon cancer. (*a–d*) Hematoxylin-eosin staining; (*e–h*) p53 immunostaining.

although his family did not meet LFS or LFL criteria (Fig. 1*a*), we screened the patient for germline *p53* mutations by sequencing analysis and a heterozygous c.584T>C mutation was found at the *p53* gene locus. The c.584T>C mutation was located in exon 6 and was associated with an amino acid substitution of threonine for isoleucine at codon 195 (Fig. 1*b*). The c.584T>C mutation has been registered as a somatic mutation on the IARC TP53 mutation database (http://www-p53.iarc.fr/),[11] and had not previously been registered as a germline mutation, indicating the c.584T>C mutation to be a novel germline mutation. The mutant allele was not detected in his parents or his brother by PCR-RFLP analysis (Fig. 1*c*), suggesting the c.584T>C mutation to be a *de novo* germline mutation.

LOH analysis using the microsatellite marker at the *p53* locus showed the presence of LOH at the *p53* locus in both glioblastoma and colon cancer (Fig. 1*d*), and in addition, further sequencing analysis revealed loss of the wild-type (wt) *p53* allele, but not the mutant allele (Fig. 1*e*). Immunohistochemical analysis showed p53 protein to be highly expressed in both glioblastoma and colon cancer, but not in the corresponding normal tissue (Fig. 1*f*). These results suggest that the Thr195-type p53 protein had accumulated in the glioblastoma and colon cancer tissues.

To evaluate the functional effect of the p.Ile195Thr mutation, we first examined the transcriptional activation potential of Thr195-type p53 by luciferase assay. p21-, BAX- and MDM2-luciferase activities were significantly lower in cells transfected with the Thr195-type p53 expression vector, Ala143-type p53 (a well-known mutant having no apparent transactivation function) vector[12] or empty vector than in the cells transfected with the wt p53 vector (Fig. 2*a*), suggesting that Thr195-type p53 lacks the transactivation function in relation to a subset of p53-responsive promoters. We next examined the ability of the Thr195-type p53 to suppress cell growth. In the G418-resistant colony formation assay, the wt p53 expression vector efficiently suppressed colony formation in comparison with an empty vector, whereas Thr195-type p53 and the Ala143-type p53 vector (a negative control)[12] did not (Fig. 2*b*), suggesting that Thr195-type p53 has no or extremely low ability to suppress cell growth. Thus, the luciferase and colony formation assays both indicated that Thr195-type p53 is functionally inactivated.

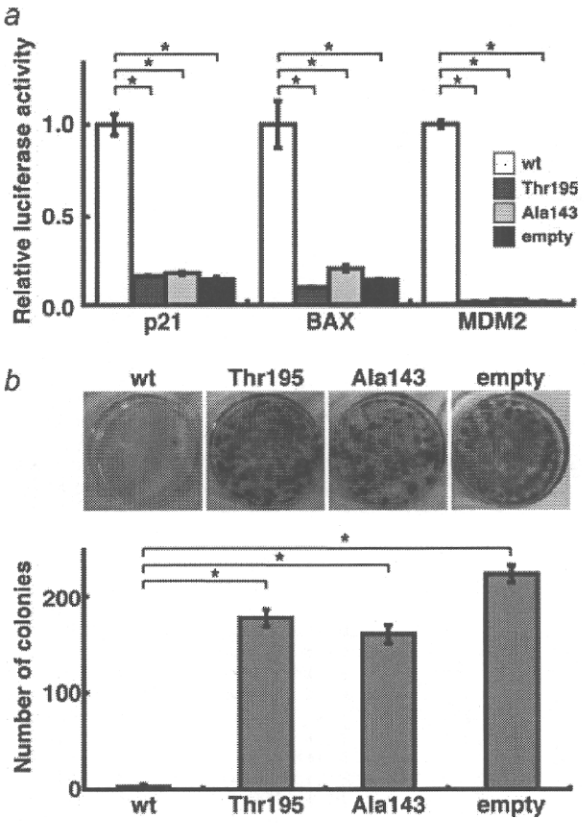

FIGURE 2 – *In vitro* characterization of Thr195-type p53 protein functions. (*a*) Evaluation of the transcriptional activation function of p53 by luciferase assay. p21-, BAX- and MDM2-luciferase activities were measured in *p53*-null H1299 cells transiently transfected with a p53 expression vector, the firefly luciferase reporter vector pGL4.10 and the transfection control vector pGL4.74. Ala143-type p53 was used as a negative control.[9] Values for luciferase activity are means ± standard deviation of 3 independent experiments. For each p53-responsive gene, the luciferase activity of cells transfected with wild-type (wt) p53 was set at 1.0. $*p < 0.01$. (*b*) Evaluation of the growth suppressive function of p53 by G418-resistant colony formation assay. Ala143-type p53 was used as a negative control.[9] Representative photographs are shown on the bar graph. Values are means ± standard deviation of 4 independent experiments. $*p < 0.0001$.

De novo germline *p53* mutations have occasionally been reported in the literature. Analyzing 268 patients with cancer in childhood with at least 1 relative affected by any cancer before 46 years of age or affected by multiple cancers for germline *p53* mutations, Chompret *et al.*[13] identified 17 mutations, 13 of which were inherited within the family whereas 4 were found to have arisen *de novo*. Analyzing 5 index cases of German or Swiss origin with cancers typical of LFS, Bendig *et al.*[14] identified 5 germline *p53* mutations, 2 of which were found to have arisen *de novo*. None of the cases having *de novo* mutations in either of these reports as well as the present study met the clinical criteria for LFS. The prevalences of *de novo* germline *p53* mutations among all germline *p53* mutations might vary depending on various factors, such as the case selection criteria for analysis and the ethnicities of the collected cases. Further collection of reports on *de novo* germline *p53* mutations may be needed to better understand this issue.

Glioblastoma and colon cancer were found in our patient. The brain tumor, but not colon cancer, is 1 of the 6 major component tumors of LFS: breast carcinoma, soft tissue sarcoma, brain tumors, bone sarcoma, adrenocortical carcinoma and leukemia. There have been reports documenting that colorectal cancers are also observed in patients with LFS, although the frequency is less than that of the 6 main LFS-associated tumors.[15,16] Nichols *et al.*[15] reported that several types of tumors, including colorectal cancer, other than the 6 main LFS-associated tumors occurred in patients with LFS at much earlier ages than expected in the general population. Furthermore, Wong *et al.*[16] showed germline *p53* mutations in 8 LFS families with early-onset colorectal cancer. Therefore, the occurrence of both glioblastoma and colon cancer in our patient is considered to be compatible with the clinical findings of patients with germline *p53* mutations.

Our patient belonged to the age group termed "adolescent and young adult" (AYA), and malignant tumors in AYA have recently been reviewed by Bleyer *et al.*[17] Some proportion of the malignant tumors in AYA are attributable to germline mutations of certain genes, such as tumor suppressor genes and oncogenes.[17] Moreover, our patient had both a glioblastoma and colon cancer. Indeed, a germline c.584T>C (p.Ile195Thr) mutation of the *p53* gene was detected in our patient and this Thr195-type p53 displayed loss of tumor suppressor activity in the regulation of transactivation and growth suppression. As, in addition to the germline mutation, loss of the remaining wt allele was seen in both tumors, p53 was biallelically inactivated in both. In conclusion, the present observations suggest germline *p53* mutations to possibly be responsible for a subset of AYA patients with multiple malignant tumors, even those not meeting the clinical criteria for LFS or LFL. Such data would contribute to the further understanding of the inheritance of susceptibility to tumor development in AYA patients.

Acknowledgements

We are grateful to Dr. B. Vogelstein for providing us with wild-type and Ala143-type p53 cDNA. We acknowledge Mr. T. Kamo for technical assistance. H.Y. is a COE research assistant in Hamamatsu University School of Medicine.

References

1. Li FP, Fraumeni JF, Jr, Mulvihill JJ, Blattner WA, Dreyfus MG, Tucker MA, Miller RW. A cancer family syndrome in twenty-four kindreds. Cancer Res 1988;48:5358–62.
2. Malkin D, Li FP, Strong LC, Fraumeni JF, Jr, Nelson CE, Kim DH, Kassel J, Gryka MA, Bischoff FZ, Tainsky MA, Friend SH. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science 1990;250:1233–8.
3. Srivastava S, Zou ZQ, Pirollo K, Blattner W, Chang EH. Germ-line transmission of a mutated p53 gene in a cancer-prone family with Li-Fraumeni syndrome. Nature 1990;348:747–9.
4. Bougeard G, Limacher JM, Martin C, Charbonnier F, Killian A, Delattre O, Longy M, Jonveaux P, Fricker JP, Stoppa-Lyonnet D, Flaman JM, Frébourg T. Detection of 11 germline inactivating TP53 mutations and absence of TP63 and HCHK2 mutations in 17 French families with Li-Fraumeni or Li-Fraumeni-like syndrome. J Med Genet 2001;38:253–7.
5. Limacher JM, Frebourg T, Natarajan-Ame S, Bergerat JP. Two metachronous tumors in the radiotherapy fields of a patient with Li-Fraumeni syndrome. Int J Cancer 2001;96:238–42.
6. Rutherford J, Chu CE, Duddy PM, Charlton RS, Chumas P, Taylor GR, Lu X, Barnes DM, Camplejohn RS. Investigations on a clinically and functionally unusual and novel germline p53 mutation. Br J Cancer 2002;86:1592–6.
7. Yamada H, Shinmura K, Okudela K, Goto M, Suzuki M, Kuriki K, Tsuneyoshi T, Sugimura H. Identification and characterization of a novel germ line p53 mutation in familial gastric cancer in the Japanese population. Carcinogenesis 2007;28:2013–18.
8. Shinmura K, Kageyama S, Tao H, Bunai T, Suzuki M, Kamo T, Takamochi K, Suzuki K, Tanahashi M, Niwa H, Ogawa H, Sugimura H. EML4-ALK fusion transcripts, but no NPM-, TPM3-, CLTC-, ATIC-, or TFG-ALK fusion transcripts, in non-small cell lung carcinomas. Lung Cancer 2008;61:163–9.

9. Tomita N, Fukunaga M, Okamura S, Nakata K, Ohzato H, Tamura S, Sugimoto K, Aihara T, Miki H, Takatsuka Y, Matsuura N, Ishikawa H, et al. The novel germline mutation of the hMLH1 gene in a case of suspected hereditary non-polyposis colorectal cancer (HNPCC) in a patient with no family history of cancer. Jpn J Clin Oncol 2004;34:556–60.
10. Shinmura K, Iwaizumi M, Igarashi H, Nagura K, Yamada H, Suzuki M, Fukasawa K, Sugimura H. Induction of centrosome amplification and chromosome instability in p53-deficient lung cancer cells exposed to benzo[a]pyrene diol epoxide (B[a]PDE). J Pathol 2008;216:365–74.
11. Petitjean A, Mathe E, Kato S, Ishioka C, Tavtigian SV, Hainaut P, Olivier M. Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database. Hum Mutat 2007;28:622–9.
12. Zhang W, Guo XY, Hu GY, Liu WB, Shay JW, Deisseroth AB. A temperature-sensitive mutant of human p53. EMBO J 1994;13:2535–44.
13. Chompret A, Brugières L, Ronsin M, Gardes M, Dessarps-Freichey F, Abel A, Hua D, Ligot L, Dondon MG, Bressac-de Paillerets B, Frébourg T, Lemerle J, et al. P53 germline mutations in childhood cancers and cancer risk for carrier individuals. Br J Cancer 2000;82:1932–7.
14. Bendig I, Mohr N, Kramer F, Weber BH. Identification of novel *TP53* mutations in familial and sporadic cancer cases of German and Swiss origin. Cancer Genet Cytogenet 2004;154:22–6.
15. Nichols KE, Malkin D, Garber JE, Fraumeni JF, Jr, Li FP. Germ-line *p53* mutations predispose to a wide spectrum of early-onset cancers. Cancer Epidemiol Biomarkers Prev 2001;10:83–7.
16. Wong P, Verselis SJ, Garber JE, Schneider K, DiGianni L, Stockwell DH, Li FP, Syngal S. Prevalence of early onset colorectal cancer in 397 patients with classic Li-Fraumeni syndrome. Gastroenterology 2006;130:73–9.
17. Bleyer A, Barr R, Hayes-Lattin B, Thomas D, Ellis C, Anderson B; Biology and Clinical Trials Subgroups of the US National Cancer Institute Progress Review Group in Adolescent and Young Adult Oncology. The distinctive biology of cancer in adolescents and young adults. Nat Rev Cancer 2008;8:288–98.

Biosensors and Bioelectronics 25 (2009) 22–27

Contents lists available at ScienceDirect

Biosensors and Bioelectronics

journal homepage: www.elsevier.com/locate/bios

Development of a method to evaluate caspase-3 activity in a single cell using a nanoneedle and a fluorescent probe

Takanori Kihara[a,b], Chikashi Nakamura[c,d,*], Miho Suzuki[e], Sung-Woong Han[c], Kyoko Fukazawa[f], Kazuhiko Ishihara[a,b,f], Jun Miyake[a,b,c,d]

[a] The Center for NanoBio Integration, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
[b] Department of Bioengineering, School of Engineering, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
[c] Research Institute of Cell Engineering, National Institute of Advanced Industrial Science and Technology, Central 4, 1-1-1 Higashi, Tsukuba, Ibaraki 305-8562, Japan
[d] Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, 2-24-16, Naka-cho, Koganei, Tokyo 184-8588, Japan
[e] Department of Functional Materials Science, Faculty of Engineering, Saitama Univeristy, 255 Shimo-okubo, Saitama, Saitama, 338-8570, Japan
[f] Department of Materials Engineering, School of Engineering, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

ARTICLE INFO

Article history:
Received 25 March 2009
Received in revised form 22 May 2009
Accepted 28 May 2009
Available online 6 June 2009

Keywords:
Nanoneedle
AFM
Single cell
FRET
Caspase-3 activity
MOMENT

ABSTRACT

A method to detect an enzymatic reaction in a single living cell using an atomic force microscope equipped with an ultra-thin needle (a nanoneedle) and a fluorescent probe molecule was developed. The nanoneedle enables the low-invasive delivery of molecules attached onto its surface directly into a single cell. We hypothesized that an enzymatic reaction in a cell could be profiled by monitoring a probe immobilized on a nanoneedle introduced into the cell. In this study, a new probe substrate (NHGcas546) for caspase-3 activity based on fluorescent resonance energy transfer (FRET) was constructed and fixed on a nanoneedle. The NHGcas546-modified nanoneedle was inserted into apoptotic cells, in which caspase-3 is activated after apoptosis induction, and a change in the emission spectrum of the immobilized probe could be observed on the surface of the nanoneedle. Thus, we have developed a successful practical method for detecting a biological phenomenon in a single cell. We call the method MOlecular MEter with Nanoneedle Technology (MOMENT).

1. Introduction

The direct manipulation and delivery of molecules into a single living cell would provide a novel technology that could be useful in biology, biotechnology, and medicine (Lamontagne et al., 2008; Voldman, 2006). There are two main applications of such molecular manipulation techniques: One is to regulate cell states (e.g., cell differentiation, cell reprogramming), and the other is to reveal the state or the type of a cell; the latter can be currently accomplished by investigating the operation of biological activities by analysing gene expression levels in the cell. For the advancement of these applications, molecular handling techniques with more spatial and temporal precision are required.

We have developed a novel molecular handling technique for a single cell using the atomic force microscope (AFM), which accurately measures the displacement and force exerted on a cantilever, and an ultra-thin needle (nanoneedle) to penetrate the cell membrane with minimal cell damage. Penetration with nanoneedles smaller than 400 nm in diameter does not cause lethal damage to the plasma membrane; because of their low invasiveness, the needles can be maintained inside the cell for more than 1 h for analysis (Han et al., 2005). Insertion of a nanoneedle through the plasma membrane can be monitored by determining the force exerted on the nanoneedle (Obataya et al., 2005c). This procedure allows us to deliver molecules that are attached onto the nanoneedle surface into the cell (Han et al., 2005; Obataya et al., 2005a).

Many kinds of fluorescent probes have been used to detect enzymatic activity (Miyawaki, 2003; Tsien, 2005). The sensing mechanism of such probes is usually based on fluorescent resonance energy transfer (FRET) between green fluorescent protein (GFP) and its colour variants for applications in the fields of cell biology, molecular biology, and developmental biology (Miyawaki, 2003; Pollok and Heim, 1999). In many cases, this system requires a construct such that an enzyme sensitive sequence is placed between two fluorescent proteins. Recently, a FRET-based system using an engineered GFP with a chemical FRET partner was reported. This chimeric FRET-based system aims to place the chemical fluorophore considerably close to and suitably oriented to GFP for efficient FRET using a site-directed mutagenesis approach (Suzuki et al., 2003, 2004).

* Corresponding author at: Research Institute of Cell Engineering, National Institute of Advanced Industrial Science and Technology, Central 4, 1-1-1 Higashi, Tsukuba, Ibaraki 305-8562, Japan. Tel.: +81 29 861 2445; fax: +81 29 861 3049.
E-mail address: chikashi-nakamura@aist.go.jp (C. Nakamura).

0956-5663/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.bios.2009.05.036

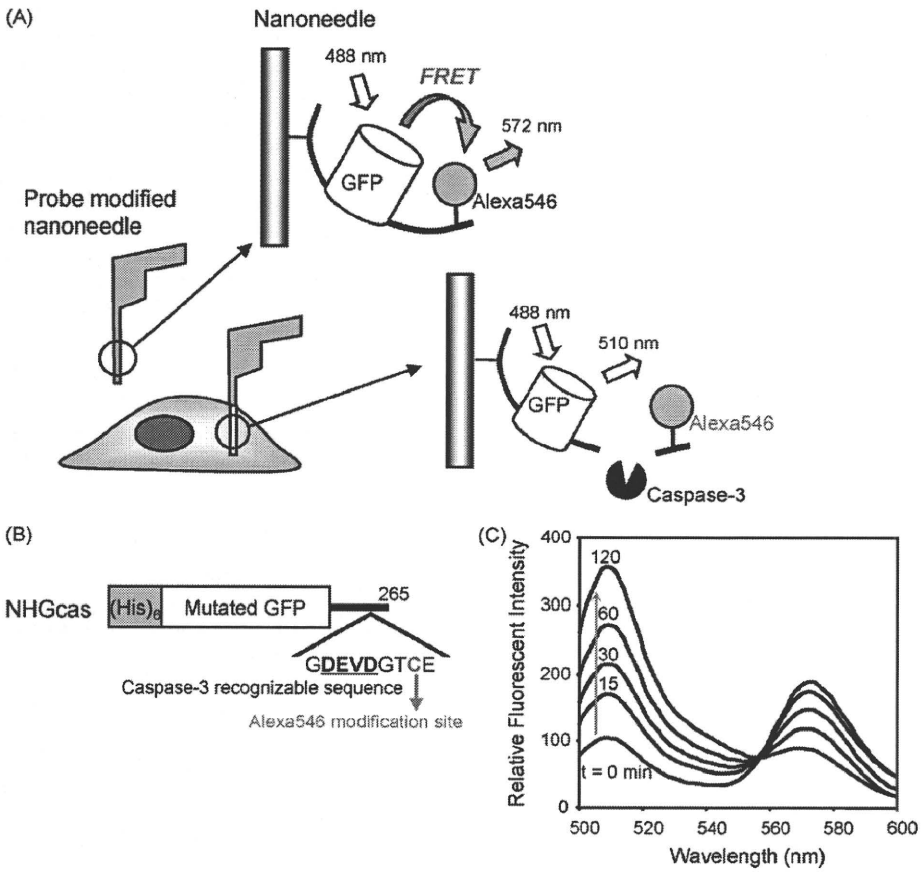

Fig. 1. Structure of the NHGcas546 probe. (A) Schematic diagram of the new biosensor for intracellular enzyme assay. A nanoneedle carrying immobilized probe molecules is inserted into a cell and the probe on the nanoneedle reacts with the enzyme inside the cell. The immobilized NHGcas546 probe emits at 572 nm upon excitation at 488 nm; after digestion with caspase-3 in an apoptotic cell, the immobilized NHGcas546 emits at 510 nm. (B) Primary structure of NHGcas. NHGcas was constructed using mutated GFP, a caspase-3 recognizable sequence (DEVD), and a Cys modified with Alexa Fluor 546. NHGcas has a $(His)_6$ tag at its N-terminus. (C) Emission spectra of NHGcas546 during digestion with recombinant caspase-3. NHGcas546 was incubated with caspase-3 for 120 min in solution. Excitation was at 488 nm. The fluorescence emission spectrum of the probe changed in caspase-3 solution and the intramolecular FRET signal was abolished in a time-dependent manner.

For the procedure, these fluorescent probes need to be expressed or introduced into the cells after fabrication by chemical or physiological methods. However, those introduction strategies have some restrictions on applicable cell types and growth conditions, and suffer from reduced viability after invasive damage. It is often difficult to employ these techniques in primary cultured cells or normal human cells. We have developed a molecular manipulation technique for a single cell that might have the potential to deliver many types of molecule into various cell types, including normal human cells, without invasive damage (Han et al., 2008). To combine this technique with the use of fluorescent chimeric probes, it is necessary to adapt the probes to function on the surface of the nanoneedle inside the cell. This combination scheme is based on the concept of a biosensor for real-time monitoring of biological phenomena in a single living cell (Fig. 1A).

In this study, a FRET-based probe (NHGcas546) functioning on a nanoneedle surface was developed to quantify the activity of caspase-3, a key enzyme of apoptosis. NHGcas546 consists of an engineered GFP with an N-terminal $(His)_6$ tag, a sequence (Asp-Glu-Val-Asp) recognizable by caspase-3 (Thornberry et al., 1997), and a cysteine for site-directed modification with an Alexa Fluor 546 dye at its C-terminus (Fig. 1B). The NHGcas546 was attached to the surface of the silicon cantilever by chelate bonding of the $(His)_6$ tag. This probe on an AFM cantilever was demonstrated to change its FRET signal as a result of cleavage by recombinant caspase-3. The NHGcas546 immobilized on a nanoneedle was then inserted into a single apoptotic cell assumed to express caspase-3 and a change in the emission spectrum was observed. Thus, the development of a new device like a micro-biosensor to investigate enzymatic activity in a single cell has been accomplished. We named the new method MOlecular MEter with Nanoneedle Technology (MOMENT). MOMENT allows the measurement of the activity of a desired molecule in a cell at that moment.

2. Materials and methods

2.1. Materials

AFM silicon cantilevers (CONT and ATEC-CONT) were purchased from Nanosensors (Neuchatel, Switzerland). The BCA protein assay kit and Halt EDTA-free protease inhibitor cocktail were purchased from Pierce Biotechnology (Rockford, IL). 2-Methacryloyloxyethyl phosphorylcholine (MPC) and *p*-nitrophenyloxycarbonyl poly(ethylene glycol) methacrylate (MEONP) were synthesized using previously reported methods (Ishihara et al., 1991; Konno et al., 2004). 3-Methacryloxypropyl triethoxysilane (MTES) was purchased from Shin-Etsu Chemical Co. Ltd. (Tokyo, Japan). Alexa Fluor 546 C5 maleimide was purchased from Invitrogen (Carlsbad, CA). TALON Metal Affinity Resin and *Escherichia coli* BL21(DE3)pLysS were purchased from Takara Bio Inc. (Shiga, Japan). NAP-10 desalting columns were purchased from GE Healthcare Life Sciences (Buckinghamshire, England).

Active human caspase-3 was purchased from Medical & Biological Laboratories Co. Ltd. (Nagoya, Japan). AB-NTA was purchased from Dojindo laboratories (Kumamoto, Japan). Other reagents were purchased from Wako Pure Chemical Industries Ltd. (Osaka, Japan) or Sigma–Aldrich (St. Louis, MO).

2.2. Plasmid construction

The protein part of the probe, called NHGcas, was derived from the GFP variant UV5casS02tag (Suzuki et al., 2005). The construction detail is described in Supplementary methods. The resulting construct was named pNHGcas.

2.3. Purification of NHGcas

E. coli BL21(DE3)pLysS was transformed with pNHGcas. LB medium supplemented with 34 μg/mL chloramphenicol, 100 μg/mL ampicillin, and 0.5-mM IPTG was inoculated with transformed *E. coli* in the growth phase and cultured at 26 °C for 3 h. The cells were harvested by centrifugation at 8000 × *g* for 15 min and resuspended in PBS. The cell suspension was suspended and vortexed with lysis buffer: PBS containing 10% glycerol, 0.5% Triton-X-100, and an EDTA-free protease inhibitor cocktail. The resulting lysate was sonicated and centrifuged at 8000 × g for 15 min. The supernatant was applied to a TALON Metal Affinity resin column. After serial washing with lysis buffer and modified lysis buffer supplemented with 2-mM imidazole, the NHGcas bound to the resin through its $(His)_6$ tag was eluted with lysis buffer containing 100-mM imidazole. The eluted protein was desalted on a NAP-10 gel permeation column. The desalted NHGcas was collected to determine its concentration with the BCA kit. Purified and quantified NHGcas was analyzed by 10% SDS-PAGE. The schematic structure of NHGcas is shown in Fig. 1B.

2.4. Preparation of the caspase-3 activity sensor probe

The caspase-3 sensor probe, NHGcas546, was prepared from NHGcas by chemical modification with Alexa Fluor 546. An aliquot of purified NHGcas dissolved in PBS containing 10% glycerol, 0.5% Triton-X-100, and 1-mM DTT was kept at room temperature for 15 min; thereafter, the remaining DTT was removed using a NAP-10 gel permeation column equilibrated with PBS. An aliquot of eluate was immediately mixed with 10 equivalent amounts of Alexa Fluor 546 C5 maleimide solution and incubated at 37 °C for 3 h. Excess fluorescent dye was removed by dialysis against PBS. The concentration of prepared NHGcas546 was determined using the BCA kit.

2.5. Preparation of nanoneedle

The nanoneedle was prepared by focused ion beam (SMI3050, SII Nanotechnology Co. Ltd., Chiba, Japan) etching as described previously (Obataya et al., 2005b). The pyramidal silicon AFM tip (ATEC-CONT, spring constant is about 0.2 N/m) was fabricated to a needle shape of approximately 400 nm in diameter and 10 μm in length with a flat-ended cylindrical edge.

2.6. Detection of caspase-3 activity in a single apoptotic cell

HeLa cells were maintained in DMEM containing 10% FBS, and antibiotics (100 units/mL penicillin G, 100 μg/mL streptomycin sulphate). Apoptosis was induced with 10 μg/mL cycloheximide and 100 ng/mL TNF-α. We performed Annexin V assays to confirm the induction of apoptosis (data not shown). Apoptotic HeLa cells were manipulated using a combination of confocal laser scanning microscopy (CLSM) (IX81/FV1000, Olympus, Tokyo, Japan) and AFM (Nanowizard I, JPK Instruments AG, Berlin, Germany). The NHGcas546 immobilized nanoneedle was inserted into an apoptotic cell aimed toward the centre of the cell by manual operation. The nanoneedle was maintained inside the cell for 60 min and then retracted. The fluorescence of NHGcas546 on the nanoneedle was examined by CLSM before and after the nanoneedle insertion. The images obtained were analyzed using ImageJ software (National Institutes of Health, Bethesda, MD, USA). The Young's modulus of the cell was calculated in accordance with the Hertz model (Obataya et al., 2005c). The force–distance curve (z-scan speed of 6 μm/s) at the region of indentation was fitted to the Hertz model. The following equation is used in the Hertz model of indentation for cylindrical material:

$$F = \frac{2aE}{(1-\nu^2)}I,$$

where F = force, I = depth of indentation, a = radius of nanoneedle, ν = Poisson's ratio (0.5), and E = Young's modulus.

All other methods used in this work are described in Supplementary methods.

3. Results and discussion

3.1. Design of a FRET-based probe for measuring caspase-3 activity on a silicon surface

A highly sensitive fluorescent probe that can function on a nanoneedle should be a powerful tool for *in situ* characterization of enzymatic activity in a single living cell. Since the nanoneedle was prepared from a single silicon AFM cantilever, the fluorescent probe should have a solid surface binding site and an enzyme-sensitive site. Thus, a new FRET-based probe composed of fluorescent protein and dye to measure caspase-3 activity was devised in this study (Fig. 1A).

This probe was attached to the nanoneedle via the N-terminus of the protein part to detect the activity of caspase-3 as a change in the emission spectrum of the probe; the protease sensitive site was located at the C-terminus of the protein part (Fig. 1B). We used GFP as the protein part and Alexa Fluor 546 as the fluorescent dye, as shown in Fig. 1B. This probe contained the DEVD sequence between GFP and Alexa Fluor 546, which is mainly cleaved by caspase-3 (Thornberry et al., 1997).

Fig. 1C shows the changes in emission profile of NHGcas546 digested with caspase-3 in solution. The emission around 510 nm from GFP clearly increased and the emission around 570 nm from Alexa Fluor 546 simultaneously decreased in a time-dependent manner (Fig. 1C). This suggests that NHGcas546 was cleaved by caspase-3 to abolish intramolecular FRET. The ratio of the fluorescence intensities at 572 and 510 nm was 1.83 at 0 min and 0.247 after 120 min after caspase-3 action. This gave a 7.4-fold alteration of the ratio by caspase-3 digestion. The alteration of the corresponding ratio on UV5casS22tag modified with Alexa Fluor 532 was 6.4 (Suzuki et al., 2005). Alexa Fluor 532 can be weakly excited at 488 nm, producing an unfavourable background for this analysis, but Alexa Fluor 546 is scarcely excited at this wavelength. Thus, NHGcas546 is a very effective probe of caspase-3 activity for further study.

3.2. Modification of a silicon surface with NHGcas546

Many points must be considered when building a probe on a nanoneedle surface to function in the interior of a cell. One consideration is the activity of the molecular probe on the nanoneedle surface for effectively sensing enzymatic activity. Another consideration is nonspecific adsorption onto the nanoneedle surface that would be an obstacle to molecular sensing. To overcome these

issues, a MPC polymer was used to modify the silicon surface (Kihara et al., 2007). The surface of MPC polymers has abundant free water molecules (Ishihara et al., 1998; Kitano et al., 2003). Thus, proteins are exposed to circumstances similar to a solution, meaning that their conformation does not change on the MPC polymer surface. Moreover, the surface of the MPC polymer has low protein adsorption. MPC is often used as a surface-modification polymer for various materials that come in contact with tissue or blood (Goda et al., 2006; Moro et al., 2004). For this study, because silicon surface bonding groups and active ester groups for immobilizing biomolecules were required in the polymer chain, MPC, MTES, and MEONP were copolymerized (Supplementary Fig. 1). MTES is a monomer for silane coupling and MEONP is a monomer with a reactive ester group (Konno et al., 2004).

First, the silicon AFM cantilever was modified with the MPC polymer by silane coupling. Thereafter, the *p*-nitrophenyl ester group was reacted with the amino group of AB-NTA to produce NTA groups on the surface. NTA is one of the most useful matrices for immobilized metal-affinity chromatography. A Ni^{2+} was chelated to the NTA group. The Ni^{2+}-NTA group can tightly bind a $(His)_6$ tagged protein. NHGcas546 was then supposed to attach to the Ni^{2+}-NTA group on the silicon AFM cantilever.

The silicon cantilever modified with NHGcas was observed by CLSM, and successful binding of the protein was confirmed (Supplementary Fig. 2). The amount of attached NHGcas was about 3×10^4 molecules/μm^2, or approximately 3 per 10-nm square. The diameter of the GFP barrel structure is about 24 Å (Ormo et al., 1996; Yang et al., 1996). Thus, it is estimated there could be approximately 20 closely packed molecules in each 10-nm square; the occupied fraction of NHGcas was about 15%. The surface area of the nanoneedle was 13 μm^2 at 400 nm diameter. Hence, about 4×10^5 molecules of NHGcas546 could be on the surface of the nanoneedle.

3.3. Cleavage of NHGcas546 with caspase-3 on the silicon surface

To evaluate the performance of the attached NHGcas546 probe on the silicon surface, the NHGcas546-modified silicon cantilever was treated with recombinant caspase-3 and was then observed by CLSM (LSM510-META) at 488 nm laser excitation and 511–599 nm emission.

The emission spectrum of NHGcas546 on the cantilever surface showed two distinct peaks: one from fluorescent protein and one from Alexa Fluor dye (Fig. 2a). The NHGcas546 on the surface of the cantilever clearly showed an intramolecular FRET spectrum pattern. When the probe was incubated with buffer alone, the emission spectrum of the NHGcas546 probe did not demonstrate significant change. However, the intensity of the fluorescence gradually decreased by sequential observation every 5 min (Fig. 2a). This suggests that the structure of NHGcas546 might be maintained on the MPC polymer-modified solid surface but its fluorescence would be bleached by continuous excitation.

Upon treatment with 1 unit of caspase-3 at room temperature, the emission of the NHGcas546 on the silicon surface clearly decreased around 570 nm (Fig. 2b). This indicated that the NHGcas546 was digested by caspase-3, thereby cancelling the intramolecular FRET signal.

Fig. 2c shows the time-dependent changes in the ratio of emission intensities of Alexa Fluor 546 (577–599 nm)/fluorescent protein (513–534 nm) on the silicon cantilever. The emission ratio was normalized by defining the baseline ratio at 0 min of caspase-3 treatment. The rate constant of 1 unit caspase-3 action on the NHGcas546 attached to the silicon cantilever was 5.4×10^{-4}/s. The ratio of the fluorescence intensity of Alexa Fluor 546/fluorescent protein after treatment with caspase-3 was 0.59. This gave a 1.7-fold alteration of the ratio by caspase-3 digestion.

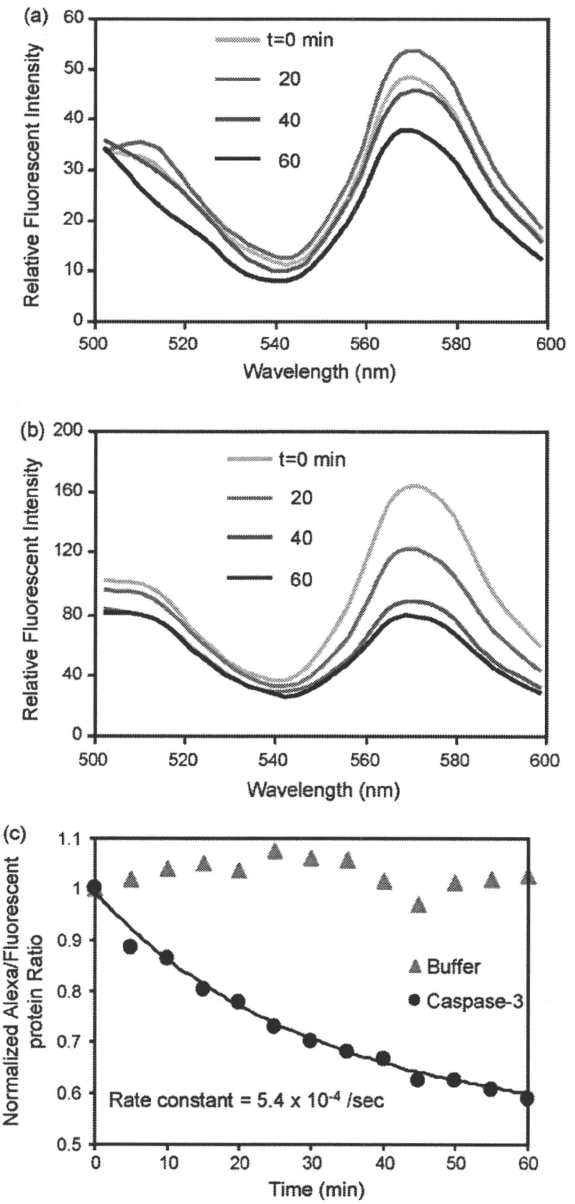

Fig. 2. Time course of caspase-3 digestion of NHGcas546 immobilized on the silicon surface. The fluorescence spectra of the immobilized NHGcas546 are overlaid (a, b). Immobilized NHGcas546 was incubated with (b) or without (a) 1 unit of recombinant caspase-3 for 60 min at room temperature on the silicon surface in 100-μL reaction solution. The fluorescence intensity of the immobilized NHGcas546 decreased in buffer solution alone (a). In contrast, the fluorescence intensity of Alexa Fluor 546 decreased in a time-dependent manner in caspase-3 solution (b). (c) First-order kinetic plot of the data from (a) and (b). The normalized ratio of the emission intensities for Alexa Fluor 546 (577–599 nm) and fluorescent protein (513–534 nm) is plotted. The points represent experimental data and the line shows the theoretical reaction kinetics determined by nonlinear least-squares regression fit to a first-order kinetic equation. Grey triangles indicate the ratio of the immobilized NHGcas546 in buffer solution, and the closed circles and black line indicate the ratio of the immobilized NHGcas546 treated with caspase-3 on the silicon surface.

The alteration of the ratio by caspase-3 digestion was larger when the NHGcas546 was free in solution (Fig. 1C, 7.4-fold) than when it was immobilized on the solid surface (Fig. 2c, 1.7-fold). One possible reason may be dysfunction of the NHGcas546 on the silicon surface caused by some denaturation during binding, which would decrease donor protein emission and consequently the efficiency of intramolecular FRET. A second possible reason may be

bleaching of the donor fluorescent protein by long-term continuous laser radiation of the target held on the solid surface, which could have a different effect than irradiation of the target in solution by the same laser. If we knew the critical point to detect some enzymatic reaction in the cell cycle or after some stimulation of the cell under study, we could avoid repeated visualization of the fluorescent probe. A third possible reason might be restricted access of caspase-3 to the immobilized NHGcas546 substrate by steric hindrance.

Next, we investigated the effect of caspase-3 concentration on the digestion of NHGcas546 on the silicon surface by changes in the ratio of donor and acceptor fluorescence emission as a routine indicator. The NHGcas546-modified silicon cantilever was treated with various amounts of caspase-3 in the reaction buffer for 30 min at 30 °C and observed by CLSM (LSM510-META) at 488 nm excitation and 513–599 nm emission. The normalized ratios of emission intensity of Alexa Fluor 546 (577–599 nm) to fluorescent protein (513–534 nm) were plotted against caspase-3 concentration (Fig. 3). The results suggested that the decay in the ratio of donor and acceptor emission intensities in response to enzyme concentration would be an adequate sensing system.

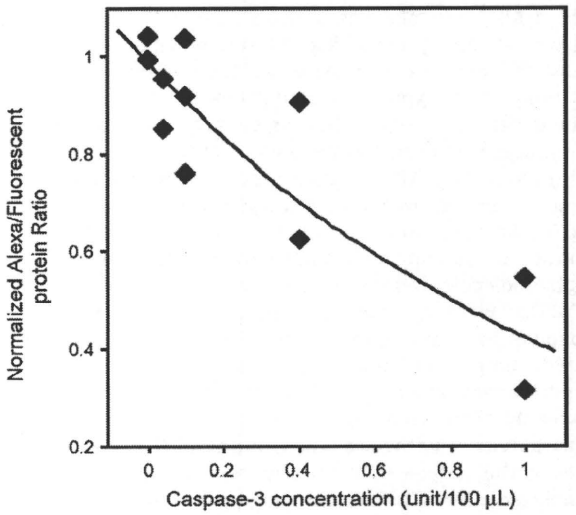

Fig. 3. Concentration-dependent caspase-3 digestion of immobilized NHGcas546 on the silicon surface. Immobilized NHGcas546 was treated with various amounts of caspase-3 for 30 min at 30 °C in 100-μL reaction solution. The normalized ratios of the emission intensities of Alexa Fluor 546 (577–599 nm) and fluorescent protein (513–534 nm) are plotted. Each mark indicates an individual experiment, and two or three experiments were performed for each caspase concentration. The line is the theoretical exponential curve derived by non-linear least-square regression fit to the plots.

3.4. Detection of caspase-3 activity in a single apoptotic cell

HeLa cells were treated with cycloheximide and TNF-α to induce apoptosis. After 2–3 h, the shape of the treated cells changed from oval to round, and apoptosis was initiated in the cells. The apoptotic cells were observed using CLSM (IX81/FV-1000) and the nanoneedle was moved toward an apoptotic cell using AFM (Fig. 4A).

Penetration of a cell with the nanoneedle can be monitored by observing a decrease or relaxation of the exerted force. Just after the nanoneedle comes in contact with the cell surface, the force on the nanoneedle increases due to resistance from indentation of the cell membrane and then suddenly decreases, indicating passage through the cell surface (Obataya et al., 2005c). After the nanoneedle travelled several μm, the force exerted on the NHGcas546-modified nanoneedle slightly increased and then flattened in the apoptotic cell (Fig. 4B). The force decreased further by about 1 nN and then flattened again. This result indicates that penetration of the probe-modified nanoneedle into an apoptotic cell was similar to penetration of a normal cell and was successfully done. The Young's modulus of the apoptotic cell was about 3.6 kPa; this value was less than the Young's modulus of normal HeLa cells (about 10 kPa). The probability of the nanoneedle being inserted

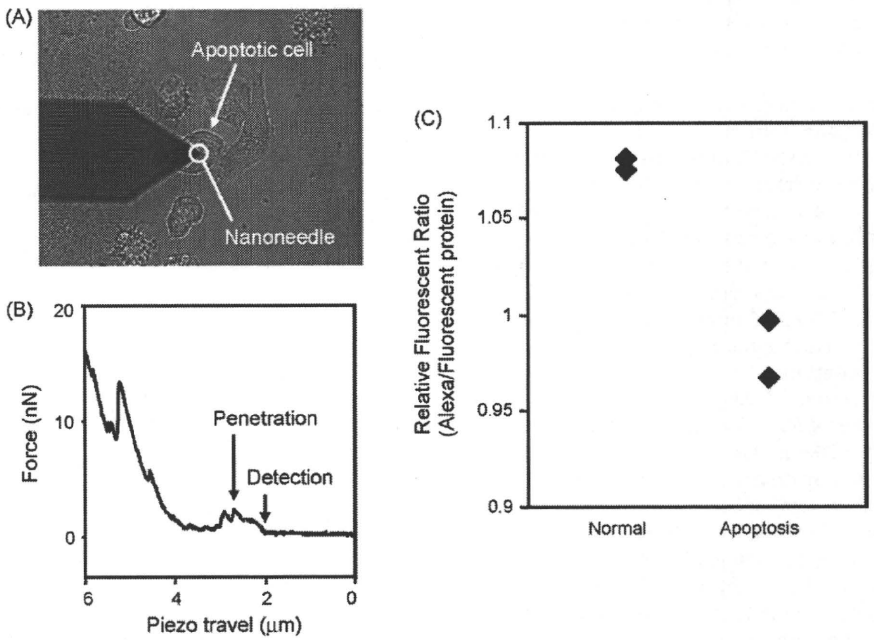

Fig. 4. Detection of caspase-3 activity in a single apoptotic cell. (A) Transmitted light image of cell manipulation using AFM. A round apoptotic cell was manipulated with the nanoneedle. (B) Force–distance curve of the nanoneedle during insertion into an apoptotic cell. The mechanical detection point and the point of penetration of the cell membrane by the nanoneedle are indicated. (C) The fluorescence ratio of NHGcas546 immobilized on the nanoneedle after insertion of the nanoneedle into a normal or apoptotic cell relative to the ratio before insertion. The NHGcas546-modified nanoneedle was maintained in the cell for 60 min at room temperature. The ratios of the normalized emission intensities of Alexa Fluor 546 (575–620 nm) and fluorescent protein (490–540 nm) are plotted.

into an apoptotic HeLa cell was over 90%. Thus, although the apoptotic cells became elastic, the probability of nanoneedle insertion was not affected by apoptosis.

To evaluate the activity of caspase-3 in apoptotic cells, the NHGcas546-modified nanoneedle was maintained inside an apoptotic cell for 60 min and the fluorescence ratios of the NHGcas546 on the nanoneedle were compared before and after insertion. The insertion of a nanoneedle with 400 nm diameter for more than 60 min does not induce lethal damage to the plasma membrane (Han et al., 2005). The fluorescence ratio decreased when the nanoneedle was inserted into apoptotic cells, but not normal cells (Fig. 4C). This result indicates successful detection of caspase-3 activation by apoptosis in a single cell using NHGcas546 immobilized on a nanoneedle.

Although we have succeeded in detecting the difference between normal and apoptotic cells by changes in the ratio, the decline of the ratio was smaller than in a previous study using chemically derived Alexa Fluor 532-modified UV5casS22tag incorporated into a single apoptotic HeLa cell, where they obtained a value which would be comparable to 0.6 in Fig. 4C (Suzuki et al., 2005). We need to consider several disadvantages of our immobilized probe system. As described in the previous section, the action of caspase-3 on the NHGcas546 probe free in solution produced a larger change in the fluorescence ratio than treatment of the same probe fixed on the silicon surface. In this study, a protein-based FRET probe was used because such probes could be highly applicable to detect many types of biological phenomena, including proteinase, kinase, or phosphatase action and changes in ion concentrations, pH, etc (Miyawaki, 2003; Sakaue-Sawano et al., 2008; Sato, 2006; Tsien, 2005). We will thus establish immobilization procedures that do not denature a protein-based probe and overcome bleaching of probes on the nanoneedle caused by continuous excitation.

The most commonly used methods for detecting apoptotic cells are the Annexin V assay and the TUNEL (TdT-mediated dUTP-biotin Nick End Labeling) assay. These methods can be used to detect apoptosis but not to continuously trace a single cell over time. FRET-based probes specific to caspase have been developed for detecting specific signal transduction during apoptosis (Takemoto et al., 2003; Tyas et al., 2000). These probes allow real-time monitoring of the enzyme activity in a single cell. However, the methods by which they are introduced into cells may cause invasive damage to the cells and cannot be applied to all cell types. Our newly devised MOMENT method involving the use of a micro-biosensor can be used for single cell analysis. Further, it can be used to specifically detect many types of biomolecules within various types of cells without causing invasive damage to the cells. In the future, we will refine this method so that it applications can be extended to many types of biomolecules.

4. Conclusions

As the demand for precise single cell manipulation to examine sequences of signal transduction increases, AFM has been intensively exploited as powerful tool in cellular and molecular biology (Lamontagne et al., 2008). In particular, AFM has been proposed to be useful for stimulation and for monitoring cellular homeostasis by introducing or extracting specific molecular entities from the cytoplasm of individual cells (Chen et al., 2007; Cuerrier et al., 2007; Han et al., 2005, 2008; Osada et al., 2003; Uehara et al., 2004). An AFM equipped with a nanoneedle would be a practical device to minimize the invasiveness of cell manipulation. This technique combined with fluorescent probes has great potential to reveal cellular states through investigating various intracellular phenomena. We have named this approach 'MOMENT'. To assess the feasibility of this approach, a new fluorescent probe for caspase-3 activity suitable for the AFM system was developed and used to detect caspase-3 activity induced in a single cell model system. Successful detection of enzymatic activity according to the cellular state suggested that this combination could be expanded to monitor many types of intracellular phenomena.

Acknowledgements

This study was partially supported by grants from the Ministry of Education, Culture, Sports, Science and Technology of Japan (Research and Development in a New Converting Field Based on Nanotechnology and Materials Sciences to JM and 19770163 to TK).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bios.2009.05.036.

References

Chen, X., Kis, A., Zettl, A., Bertozzi, C.R., 2007. Proc. Natl Acad. Sci. U.S.A. 104 (20), 8218–8222.
Cuerrier, C.M., Lebel, R., Grandbois, M., 2007. Biochem. Biophys. Res. Commun. 355 (3), 632–636.
Goda, T., Konno, T., Takai, M., Moro, T., Ishihara, K., 2006. Biomaterials 27 (30), 5151–5160.
Han, S., Nakamura, C., Obataya, I., Nakamura, N., Miyake, J., 2005. Biochem. Biophys. Res. Commun. 332 (3), 633–639.
Han, S.W., Nakamura, C., Kotobuki, N., Obataya, I., Ohgushi, H., Nagamune, T., Miyake, J., 2008. Nanomedicine 4 (3), 215–225.
Ishihara, K., Nomura, H., Mihara, T., Kurita, K., Iwasaki, Y., Nakabayashi, N., 1998. J. Biomed. Mater. Res. 39 (2), 323–330.
Ishihara, K., Ziats, N.P., Tierney, B.P., Nakabayashi, N., Anderson, J.M., 1991. J. Biomed. Mater. Res. 25 (11), 1397–1407.
Kihara, T., Yoshida, N., Mieda, S., Fukazawa, K., Nakamura, C., Ishihara, K., Miyake, J., 2007. NanoBiotechnology 3 (2), 127–134.
Kitano, H., Imai, M., Mori, T., Gemmei-Ide, M., Yokoyama, Y., Ishihara, K., 2003. Langmuir 19 (24), 10260–10266.
Konno, T., Watanabe, J., Ishihara, K., 2004. Biomacromolecules 5 (2), 342–347.
Lamontagne, C.A., Cuerrier, C.M., Grandbois, M., 2008. Pflugers Arch. 456 (1), 61–70.
Miyawaki, A., 2003. Dev. Cell 4 (3), 295–305.
Moro, T., Takatori, Y., Ishihara, K., Konno, T., Takigawa, Y., Matsushita, T., Chung, U.I., Nakamura, K., Kawaguchi, H., 2004. Nat. Mater. 3 (11), 829–836.
Obataya, I., Nakamura, C., Han, S., Nakamura, N., Miyake, J., 2005a. NanoBiotechnology 1 (4), 347–352.
Obataya, I., Nakamura, C., Han, S., Nakamura, N., Miyake, J., 2005b. Biosens. Bioelectron. 20 (8), 1652–1655.
Obataya, I., Nakamura, C., Han, S., Nakamura, N., Miyake, J., 2005c. Nano Lett. 5 (1), 27–30.
Ormo, M., Cubitt, A.B., Kallio, K., Gross, L.A., Tsien, R.Y., Remington, S.J., 1996. Science 273 (5280), 1392–1395.
Osada, T., Uehara, H., Kim, H., Ikai, A., 2003. J. Nanobiotechnol. 1 (1), 2.
Pollok, B.A., Heim, R., 1999. Trends Cell Biol. 9 (2), 57–60.
Sakaue-Sawano, A., Kurokawa, H., Morimura, T., Hanyu, A., Hama, H., Osawa, H., Kashiwagi, S., Fukami, K., Miyata, T., Miyoshi, H., Imamura, T., Ogawa, M., Masai, H., Miyawaki, A., 2008. Cell 132 (3), 487–498.
Sato, M., 2006. Anal. Bioanal. Chem. 386 (3), 435–443.
Suzuki, M., Ito, Y., Sakata, I., Sakai, T., Husimi, Y., Douglas, K.T., 2005. Biochem. Biophys. Res. Commun. 330 (2), 454–460.
Suzuki, M., Ito, Y., Savage, H.E., Husimi, Y., Douglas, K.T., 2003. Chem. Lett. 32 (3), 306–307.
Suzuki, M., Ito, Y., Savage, H.E., Husimi, Y., Douglas, K.T., 2004. Biochim. Biophys. Acta 1679 (3), 222–229.
Takemoto, K., Nagai, T., Miyawaki, A., Miura, M., 2003. J. Cell Biol. 160 (2), 235–243.
Thornberry, N.A., Rano, T.A., Peterson, E.P., Rasper, D.M., Timkey, T., Garcia-Calvo, M., Houtzager, V.M., Nordstrom, P.A., Roy, S., Vaillancourt, J.P., Chapman, K.T., Nicholson, D.W., 1997. J. Biol. Chem. 272 (29), 17907–17911.
Tsien, R.Y., 2005. FEBS Lett. 579 (4), 927–932.
Tyas, L., Brophy, V.A., Pope, A., Rivett, A.J., Tavare, J.M., 2000. EMBO Rep. 1 (3), 266–270.
Uehara, H., Osada, T., Ikai, A., 2004. Ultramicroscopy 100 (3-4), 197–201.
Voldman, J., 2006. Curr. Opin. Biotechnol. 17 (5), 532–537.
Yang, F., Moss, L.G., Phillips Jr., G.N., 1996. Nat. Biotechnol. 14 (10), 1246–1251.

Journal of Bioscience and Bioengineering
VOL. 109 No. 1, 62–66, 2010

ELSEVIER

www.elsevier.com/locate/jbiosc

Onset timing of transient gene expression depends on cell division

Kazumi Hakamada,[1,2,†] Satoshi Fujita,[1,2,†] and Jun Miyake[1,2,*]

Department of Bioengineering, School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan[1] *and Research Institute for Cell Engineering (RICE), National Institute of Advanced Industrial Science and Technology (AIST), 2-41-6 Aomi, Koto-ku, Tokyo 135-0064, Japan*[2]

Received 12 February 2009; accepted 8 July 2009
Available online 28 July 2009

By using the time lapse of both phase-contrast and fluorescent images, we examined the morphology of cells and the dynamics of gene expression (EGFP). We applied *k*-means clustering to the time course data of fluorescent intensity of EGFP and successfully found four subpopulations.

Discriminating the appropriate clusters and investigating the details of them, we found that almost all cells express the transfected gene after cell division and also found there is a strong correlation between onset timing and cell division. This result suggested that it is possible to normalize the dynamics of gene by arranging the onset timing of gene expression or by arranging the cell division.

[**Key words:** Cellular dynamics; Onset timing; Single cell analysis; Time course; Transfection mechanism; Synchronization; Gene expression]

Functions of cells such as differentiation, apoptosis, and tumorigenesis are the result of the systematic and dynamic reactions of intracellular molecules. Computational modeling, if possible, enables to predict the behavior of cells and to reproduce and design the cellular functions. Cellular dynamics, time-dependent change of interactions of intra-cellular molecules (1–6), the synthesis, degradation and interactions of intra-cellular molecules that include small molecules, DNAs, RNAs, proteins and lipids are the key themes of systems biology (7).

Cellular dynamics has conventionally been investigated by measuring the entire population of cells in the well that means superimposed value for a population of cells. Data superimposed, which are taken from the entire population of cell, do not always represent the behavior of a single cell. Even in the isogenic population, the behavior of gene expression in each cell differs from each other (8). It suggests that the measurement of gene expression of the entire population of cells as the component of cellular system has a limitation to uncover the cellular system. Although behavior of every gene is unique in each cell, they have the same behaviors against drugs (6).

We suppose there is a regulatory system to stabilize the response against the diversity of reactions (homeostasis). To understand the mechanism how the system should be organized is a key for the application of cell technologies.

Single-cell studies are the key, especially in the non-genetic variation, to elucidating the cellular dynamics. Bengtsson et al. developed the method for single-cell RT-qPCR and revealed the differences of the amount of mRNAs in the cell (9). Hirai et al. reported the monitoring of the morphology of cell and evaluating the velocity of cell spreading (10). Feinerman et al. reported the difference of protein expression between single T cells and revealed that this difference generated dispersion in responsiveness (11). Sigal et al. reported that they developed the method to extract cell cycle dependent nucleic protein by using single cell analysis (12). Although they open up the method for cellular dynamics by single cell time-course analysis, problems still remain: it is difficult to elucidate the onset timing of cellular dynamics.

Although recent advances enhanced our understanding of cellular dynamics, two questions still remain. First of all, almost all cellular reaction occurs in the cell but conventional biological methods only allow us to deal in not single cell but the entire population of cells. It means that conventional methods provide the superimposed data of cellular dynamics. Superimposed data implies that all cells in the population are of the same character and the same behaviors. Instead of the implication described above, recent study reveals cellular heterogeneity, which was caused by genetic and/or non-genetic variation (13). It suggests that the dynamics observed, which were obtained by analyzing the entire population of cells, were not always consistent with the dynamics of single cells. To elucidate cellular dynamics, it is necessary to observe a single cell and collect a vast number of them to analyze statistically. Secondly, because almost all reactions in the cell occur sequentially, the onset timing of each reaction is a key to understand target reaction. Furthermore, onset timing of these reactions is different in each cell. Considering the above two points, we anticipate that it is possible to observe intrinsic

* Corresponding author. Department of Bioengineering, School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan. Tel./fax: +81 3 5841 1791.
E-mail address: jun-miyake@will.dpc.u-tokyo.ac.jp (J. Miyake).
† These authors contributed equally to this work.

1389-1723/$ - see front matter © 2009, The Society for Biotechnology, Japan. All rights reserved.
doi:10.1016/j.jbiosc.2009.07.003

aspects of cellular dynamics and to discriminate between cellular states without receiving influences from genetic and/or non-genetic diversity. To understand cellular dynamics, it is necessary to develop a method that can measure the onset timing of gene expression in each cell.

In this study, to develop the method for measuring the onset timing of gene expression, we observed the above two issues; the time dependent morphological change of single cells and the time course of intensity of fluorescence of gene expression. We found the strong correlation between the cell division and onset timing of gene expression and showed that there is a common manner in the population of cells — that they have quite different behaviors.

MATERIALS AND METHODS

Cultivation and transfection of cells HeLa cells were obtained from commercial sources and were cultivated in an incubator at 37 °C in an atmosphere of 5% CO_2 in air. For the reporter assay, HeLa cells were seeded at 2×10^4 cells per well in 24-well culture plates. Cells were transfected with pEGFP-N1, which encodes EGFP under the control of the CMV promoter, in the presence of Lipofectamine™ LTX (Invitrogen, Carlsbad, CA). After transfection, fluorescent intensity was measured immediately with a Programmable Cellular Image Tracer, which had been co-developed with Olympus (Olympus, Tokyo, Japan). Since Tseng et al. and Cohen et al. identified a correlation between uptake of plasmid, which was controlled by CMV promoter, into individual cells and GFP expression by those cells (14, 15), we used CMV as a promoter in this study.

Data processing for *in silico* isolation and synchronization of cells After transfection with pEGFP-N1, phase-contrast and fluorescent images of cells in each well of 24-well culture plates were recorded at intervals of 15 min for two days, and their exposure times were 350 ms and 35 ms, respectively. In our long-term analysis of the fluorescent images of cells, each cell was independently recognizable and its fluorescence was measurable with our software, which had been co-developed with Olympus (Olympus, Tokyo, Japan). At first, fluorescent images were smoothed by Gaussian filter and their backgrounds were corrected by rolling ball method. In the next, images were binarized by threshold (intensity of object was at least 16 times higher than that of background), which was heuristically decided by comparing fluorescent image. The value of the threshold depends on the experimental condition and time of exposure, it should be optimized again in each experiment. We extracted the following three parameters for data preprocessing with the newly developed software:

$$\begin{aligned} &\text{Start time} = \text{Time of onset of expression of EGFP;} \\ &\text{Last time} = \text{Last time that the software recognized the cell;} \\ &\text{RT} = \text{Last time} - \text{Start time.} \end{aligned} \tag{1}$$

"Start time" was defined as the first time that the intensity of fluorescence from EGFP exceeded the optimized threshold for detection. "Last time" was defined as the first time that the software failed to recognize a cell. "RT" was calculated from the above definitions. Using these three parameters, we selected cells that satisfied the equation RT $\geq 2 \times$ interval time. Next, we standardized the intensity of fluorescence from each cell as follows:

$$I_i(t)_{\text{std}} = \frac{I_i(t) - min(I_i)}{max(I_i) - min(I_i)}, \tag{2}$$

where $I_i(t)_{\text{std}}$ is the i-th cell with the standardized intensity of fluorescence at time t. $I_i(t)$ is the i-th cell with a certain intensity of fluorescence at time t (16, 17).

The timing of cell division was determined from phase-contrast images. At the M phase, since a parent cell detached from the culture dish before dividing into two daughter cells, it was easy to distinguish dividing cells from other cells.

To extract features of the time course of changes in fluorescence due to EGFP, we applied "k-means clustering" to our time-course data and defined the center of cluster as a superimposed intensity of the fluorescence of cluster member. The silhouette value was used to evaluate the appropriate number of clusters (18). In this study, we used the Euclidean distance as the dissimilarity. The silhouette value $s(i)$ was defined as the dissimilarity between clusters. For example, take any object i in the data set, and stipulate that this object belongs to cluster A. When cluster A contains objects apart from i, we calculate

$$a(i) = \text{average dissimilarity of } i \text{ to all other objects in cluster } A. \tag{3}$$

Next, cluster C is considered, and we calculate

$$d(i, C) = \text{average dissimilarity of } i \text{ to all other objects in cluster } C. \tag{4}$$

After we have calculated $d(i, C)$ for all clusters where $C \neq A$, $b(i)$ is defined as follows:

$$b(i) = \min_{C \neq A} d(i, C). \tag{5}$$

Using $a(i)$ and $b(i)$, $s(i)$ is defined as follows:

$$s(i) = \frac{b(i) - a(i)}{max\{a(i), b(i)\}}. \tag{6}$$

The silhouette value is between -1 and 1 and, as it approaches 1, the objective data belong to a more appropriate cluster.

RESULTS AND DISCUSSION

The behavior of a population of cells and that of each individual cell in the population are quite different. To confirm the necessity of single-cell analysis, we investigated this difference in a population of HeLa cells (Fig. 1). As we expected, the time courses of changes in fluorescent intensity due to EGFP of individual cells were very different from that of the entire population of cells, even though all cells have been transfected and cultured under the same conditions. Thus, clonal cells exhibited heterogeneity and their dynamics were obviously influenced by both their cellular state and their environment, for example, adhesion to adjacent cells.

To elucidate the cellular dynamics of individual cells, we analyzed a large number of time courses of changes in fluorescence of single cells and extracted their common features. Fig. 2 shows the schematic diagram of our method. We recognized 18,570 individual cells and recorded the respected time courses. Then, we selected 9163 cells that provided data for more than 100 min. To extract the features for these cells, we performed k-means clustering (9163 cells; $k = 2$ through 10), and then we evaluated the results by calculating

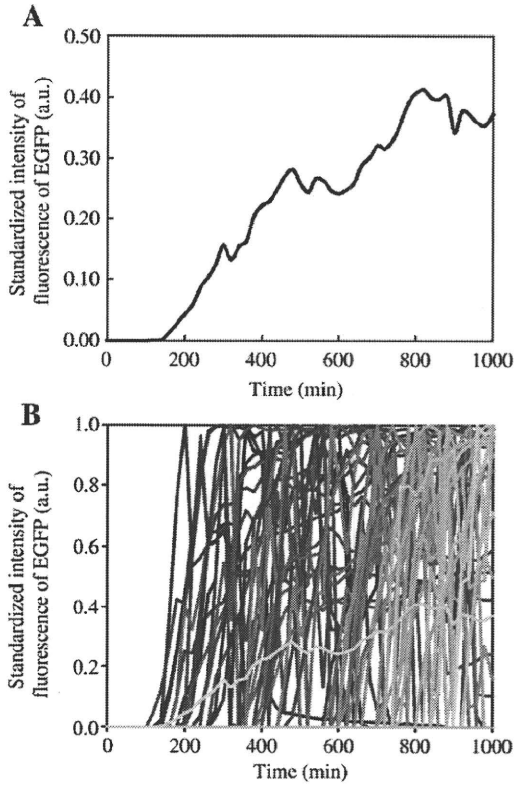

FIG. 1. (A) Superimposed intensity of fluorescence from EGFP. Superimposed intensity of fluorescence was calculated from data obtained with HeLa cells. Data were recorded for 15 min for two days. The abscissa and ordinate show the time and the standardized intensity of fluorescence from EGFP, respectively ($n = 70$ cells). (B) Time course of changes in fluorescence from single cells. Each line shows the time course of changes in the intensity of fluorescence from EGFP in a single cell. Data were calculated using Eq. (2) in Materials and methods ($n = 70$ cells).

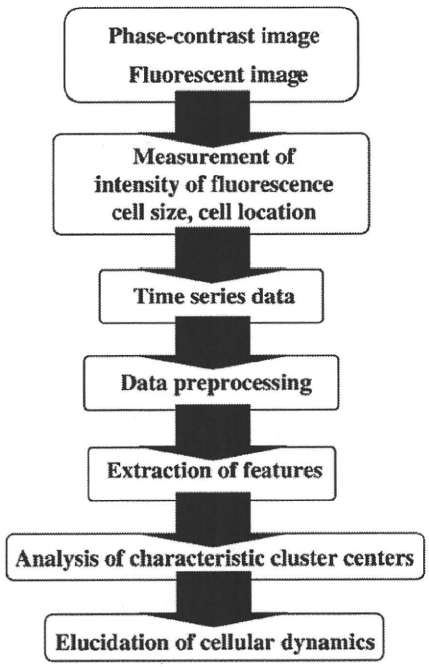

FIG. 2. Schematic diagram of data processing for *in silico* identification and extraction of the dynamics of cells.

silhouette values. Fig. 3A shows the distributions of silhouette values for each cluster (k value). The abscissa and the ordinate represent, respectively, the silhouette value and the number of cells. When the number of clusters was 4 ($k=4$), the mean of the distribution of the silhouette value had the highest value. Our analysis implied that there are at least four subpopulations in the entire population of cells that we examined, in particular, our analysis indicated that there are four characteristic time courses ($k=4$).

The profiles of four centers of cluster at various times were shown in Fig. 3B. In this study, since cells were transfected with the EGFP-expression plasmid and EGFP mRNA was transcribed from a CMV promoter, we expected that transcriptional activity would be continuous throughout the cell cycle, and the intensity of fluorescence from EGFP would increase gradually. The profiles of the center of cluster 1 and cluster 4 showed that their time course of fluorescent intensity increased gradually. Because the center of cluster 1 fluorescent intensity increased more rapidly than that of cluster 4, we focused on the center of cluster 1 and the members around the center of cluster 1 as cells with an ideal and most characteristic pattern of gene expression.

Phase-contrast and fluorescent images for members of cluster 1 concurrently in every 15 min for two days were shown in Figs. 4A and B. In the majority of cells in cluster 1, the onset timing of expression of EGFP was observed after cell division, suggesting that the timing of the onset of expression of the gene for EGFP and cell division were closely correlated.

Confirming statistically about this correlation, we recorded cell division and the onset timing of expression of EGFP (Fig. 5). In the scatter plots of the timing of the onset of gene expression versus that of cell division (Fig. 5A), linear regression analysis suggested a strong correlation between the timing of cell division and the timing of onset of gene expression (multiple correlation coefficient, $R^2=0.9238$, $P=0.01$). Fig. 5B shows distribution of the time lag between the onset timing of expression of the gene for EGFP and cell division. The distribution is unimodal, with a peak at approximately 80 min, which was the time when the intensity of fluorescence from EGFP in most cells exceeded the optimized threshold for detection of such fluorescence.

Other members of clusters ($k=4$) with different features were investigated. First, we focused on the cells around the center of cluster 2, from which the intensity of fluorescence due to EGFP remained high for 90 min. But, 53 cells (52%), whose fluorescence might have been derived from non-degraded EGFP, were dying or already dead on our initial phase-contrast images. When we monitored the fluorescence from spherical cells, which were generated as a result of cell death or cell division, the fluorescence from EGFP in each cell tended to increase. In fact, almost all spherical cells, in which the intensity of fluorescence was enhanced in this way, seemed to be dying just after the initial measurements had been made. The intensity of fluorescence due to EGFP in cells was maintained, without decay, for at least 90 min. Thus, cells in cluster 2 were seriously damaged and seemed unable to express EGFP normally.

In the case of cells in cluster 3, where the intensity of fluorescence decreased gradually, 33 cells (33%) were dead within 90 min. We were aware of a steady decrease in the fluorescence from each cell with the fragmentation of cells that occurred after they died. Thus, cells in cluster 3 were also damaged like the cells in cluster 2 and seemed unable to express EGFP normally.

The changes in the intensity of fluorescence in the case of cells in cluster 4, remained small for 90 min, but the intensity did increase gradually. When we investigated the correlation between the timing of cell division and the onset timing of expression of EGFP in this case, we obtained a high multiple correlation coefficient ($R^2=0.8454$, $P=0.01$). This result suggests that cells in both cluster 1 and cluster 4 might belong to the same biological group and had been successfully

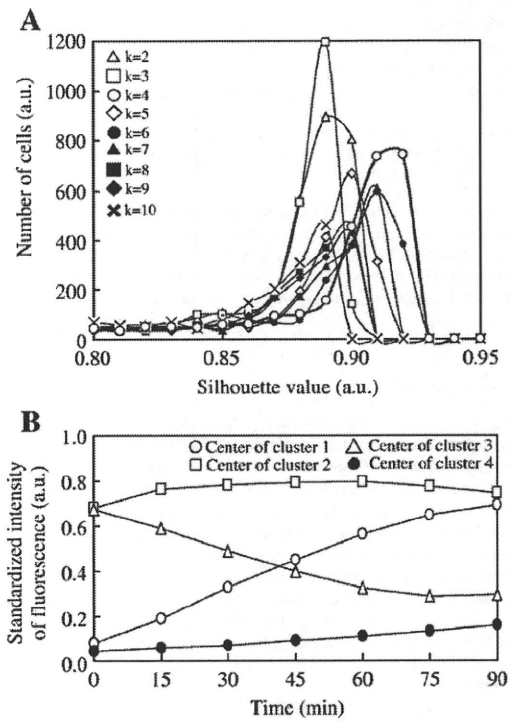

FIG. 3. (A) Distributions of silhouette values for each value of k. Each distribution of silhouette values was calculated for a specific cluster number ($k=2$ through 10). Each symbol refers to the results for a different value of k. The abscissa and ordinate represent the silhouette value and the number of cells, respectively. The mode of the silhouette value of each distribution was highest in the case of $k=4$ (thick line). (B) Time course of each center of cluster ($k=4$; see text for details). The abscissa and ordinate show the time course after transfection and standardized intensity of fluorescence from EGFP, respectively.

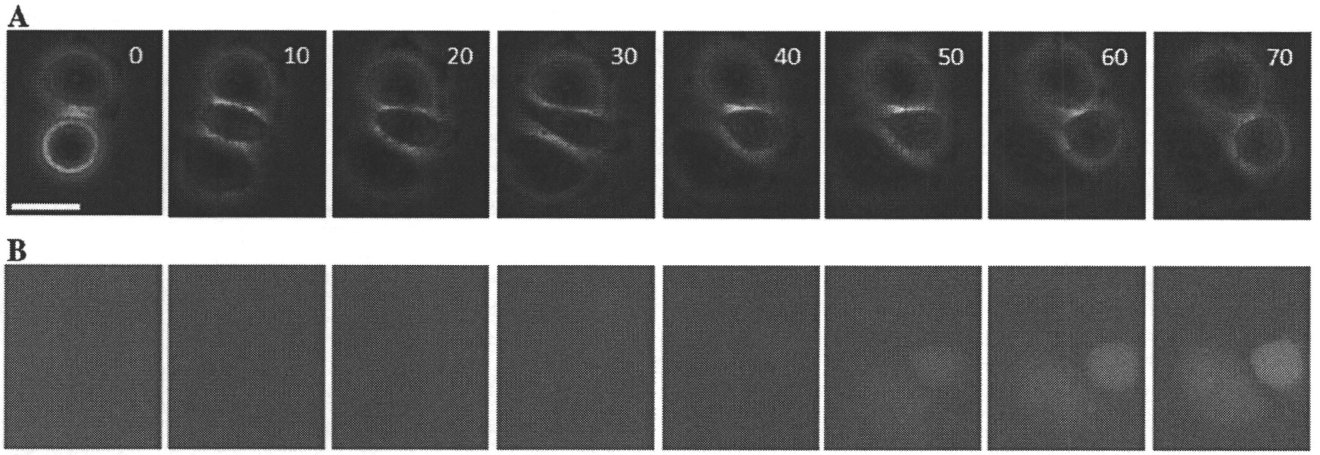

FIG. 4. Typical images of cell, which expressed EGFP after cell division. Series of sequential phase-contrast and fluorescent images of HeLa cells that had been transfected with a plasmid that encoded EGFP and images were taken immediately after transfection. These images were recorded every 15 min for 150 min and showed every 30 min. (A) Phase-contrast images and (B) fluorescent images. t = time in minutes; scale bar = 25 μm.

transfected without adverse effects, even though the intensity of fluorescence from EGFP in the cells in cluster 4 was lower than that in the cells in cluster 1. The difference between two clusters is thought that the copy number of transfected gene of the EGFP is different.

Cohen et al. and Tseng et al. reported that fluorescent intensity of transfected gene was strongly correlated with intra-cellular plasmid copy number (14, 15). On the contrary Chang et al. reported that flow cytometric analysis of a multipotent mouse haematoic cell line revealed an approximately 1000-fold range in the level of the constitutively expressed stem-cell-surface marker Sca-1 among individual cells derived from clonal cell population (19). These reports suggested that fluorescent intensity derived from transfected gene highly depends on the copy number of plasmid and their heterogeneity. It is difficult to discriminate whether the influence on the intensity of fluorescence is the difference of copy numbers of the plasmid or the heterogeneity. In this study, although we were able to discriminate the difference of the time course of fluorescent intensity by using single cell analysis, it is necessary to investigate the relationship between the copy number of plasmids and the heterogeneity to elucidate the quantitative dynamics of fluorescent intensity.

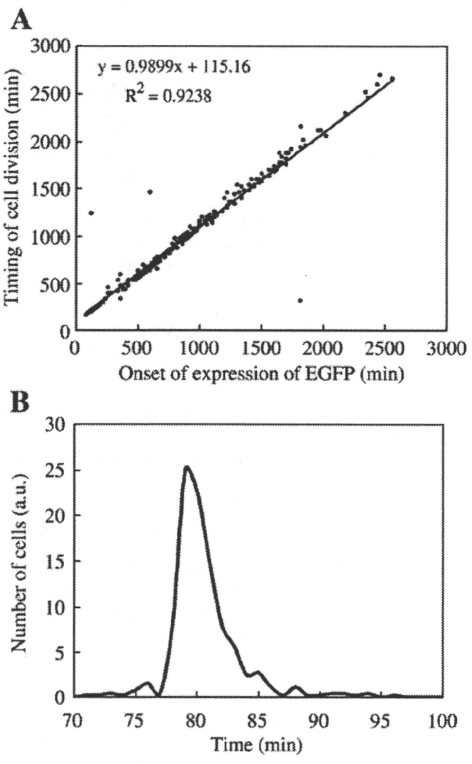

FIG. 5. (A) Correlation between the timing of cell division and the onset timing of expression of EGFP. The abscissa represents the first time that intensity of fluorescence from EGFP exceeded the optimized threshold value. The ordinate represents the timing of the start of cell division. The thick line is the regression line. The equation and R^2 represent the regression equation and the multiple correlation coefficient ($n = 221$ cells). (B) Distribution number of cells of the time lag between the onset of expression of EGFP and the start of cell division (time in minutes) ($n = 512$ cells).

The numbers of members in each cluster are 2049 (cluster 1; 22.4%), 1219 (cluster 2; 13.3%), 1047 (cluster 3; 11.4%) and 4848 (cluster 4; 52.9%), respectively ($k = 4$). The sum of the percentages of cells in cluster 1 (22.4%) and cluster 4 (52.9%) is 75.3%, corresponding to the efficiency of transfection of HeLa cells (80%, unpublished data). It seems plausible, therefore, that cells in cluster 1 and cluster 4 were transfected without adverse effects and that EGFP was expressed appropriately.

In this study, we investigated the common manner in the population of cells that each cell has quite different behaviors. Although single cells in the same population have quite different behaviors, onset timing of expression of transfected gene has a strong correlation with cell division. In the case of EGFP encoded by a plasmid, our results show that onset of expression of the transfected gene occurred after cell division. It suggests that the transfected gene was only transferred to the cytoplasm by endosomes without any movement into the nucleus, because movement of a plasmid into the nucleus is prevented by the nuclear membrane. At M phase and, in particular, at prophase, when the nuclear membrane is temporarily disassembled (20), the transfected plasmid can easily move into the nucleus. Then, when the nuclear membrane has re-formed, gene on the plasmid can be transcribed to yield mRNA. Our results reflect the previous report, by Escriou et al. (21), that the nuclear import of plasmids is greatly facilitated by mitosis and vice versa, transfection efficiency of G1 arrested cells by aphidicolin, which arrest cell cycle at S phase, was 20-fold lower than that of control asynchronous cells (22). In Fig. 5, there are few cells plotted away from the regression line and only one cell expressed the EGFP before cell division. It suggested that the mechanism of transportation into nuclear without cell

division was a minor route and was not expected to transport plasmid into nuclear before cell division. This manner means that the onset timing of gene expression is highly dependent by cell division, it is possible to discriminate the onset timing of gene expression.

In the case of non-dividing cells, because transporting gene to nuclear is difficult by using cell division, it is necessary to develop a method that can transport genes to nuclear. To transport gene to the nuclear without cell division, Miller et al. reported that 176 bp portion of the smooth muscle γ-actin (SMGA) promoter can mediate plasmid nuclear import specifically in smooth muscle cells (23). By the promotion of cytoplasm-to-nuclear recruitment and enhancement of transcription, Epstein–Barr nuclear antigen 1 (EBNA1) and oriP sequence accelerate the gene delivery and expression (24). Both methods exploit the specific sequences, which are SMGA and oriP, respectively and these sequences are bound by serum response factor (SRF) and EBNA1 then bound by importin α then transported into nuclear by exploiting nuclear pore (23,25). These reports suggested that, by using nuclear pore, plasmid can be transported independently of cell division. Our result suggested that developing the transfection method by exploiting nuclear pore might be the main target to improve transfection efficiency of non-dividing cells.

We focused on the time lag between the onset of gene expression of EGFP and cell division, as shown in Fig. 2B. Remington reported that the maturation of GFP takes 30 min (26), Tsien et al. also reported that the time constant of S65T-GFP is 0.45 h (27, 28) and, in HeLa cells, it is generally accepted that the M phase lasts about 60 min. These observations suggest that the time required for the initiation of fluorescence from EGFP after cell division is approximately, at least, 90 min, and they also support our observation of a peak around 80 min after cell division. This conclusion was also valid when we used another constitutive promoter, namely, that of SV40 (data not shown).

In this study, we found that almost all cells express the transfected genes after cell division and we showed that this manner had a strong correlation. It suggests that the dynamics of genes in each cell can be normalized by arranging the onset timing of gene expression or by arranging the cell division. By using single cell time-course analysis, we can extract an accurate time course of gene expression. Extracting and normalizing the expression of genes make it possible to understand and predict the cellular dynamics. We are currently investigating the potential of expanding our method by the various cell lines.

ACKNOWLEDGMENTS

This study was performed as part of "Analytical Technology Development of Dynamic Networks Inside Living Cells/Technology Development for the Real Time Analysis of the Network In the Living Cells", which is funded by the New Energy and Industrial Technology Development Organization (NEDO) of Japan. This work was partly supported by Grant-in-Aid for Young Scientists (B) no. 20760534.

References

1. **Ankers, J. M., Spiller, D. G., White, M. R., and Harper, C. V.:** Spatio-temporal protein dynamics in single living cells, Curr. Opin. Biotechnol., **19**, 375–380 (2008).
2. **Sigal, A., Milo, R., Cohen, A., Geva-Zatorsky, N., Klein, Y., Liron, Y., Rosenfeld, N., Danon, T., Perzov, N., and Alon, U.:** Variability and memory of protein levels in human cells, Nature, **444**, 643–646 (2006).
3. **Lahav, G., Rosenfeld, N., Sigal, A., Geva-Zatorsky, N., Levine, A. J., Elowitz, M. B., and Alon, U.:** Dynamics of the p53-Mdm2 feedback loop in individual cells, Nat. Genet., **36**, 147–150 (2004).
4. **Geva-Zatorsky, N., Rosenfeld, N., Itzkovitz, S., Milo, R., Sigal, A., Dekel, E., Yarnitzky, T., Liron, Y., Polak, P., Lahav, G. and Alon, U.:** Oscillations and variability in the p53 system, Mol. Syst. Biol., **2**, 2006 0033, (2006).
5. **Kholodenko, B. N.:** Cell-signalling dynamics in time and space, Nat. Rev. Mol. Cell. Biol., **7**, 165–176 (2006).
6. **Cohen, A. A., Geva-Zatorsky, N., Eden, E., Frenkel-Morgenstern, M., Issaeva, I., Sigal, A., Milo, R., Cohen-Saidon, C., Liron, Y., Kam, Z., Cohen, L., Danon, T., Perzov, N., and Alon, U.:** Dynamic proteomics of individual cancer cells in response to a drug, Science, **322**, 1511–1516 (2008).
7. **Kitano, H.:** Biological robustness, Nat. Rev. Genet., **5**, 826–837 (2004).
8. **Shahrezaei, V. and Swain, P. S.:** The stochastic nature of biochemical networks, Curr. Opin. Biotechnol., **19**, 369–374 (2008).
9. **Bengtsson, M., Hemberg, M., Rorsman, P., and Stahlberg, A.:** Quantification of mRNA in single cells and modelling of RT-qPCR induced noise, BMC Mol. Biol., **9**, 63 (2008).
10. **Hirai, H., Umegaki, R., Kino-Oka, M., and Taya, M.:** Characterization of cellular motions through direct observation of individual cells at early stage in anchorage-dependent culture, J. Biosci. Bioeng., **94**, 351–356 (2002).
11. **Feinerman, O., Veiga, J., Dorfman, J. R., Germain, R. N., and Altan-Bonnet, G.:** Variability and robustness in T cell activation from regulated heterogeneity in protein levels, Science, **321**, 1081–1084 (2008).
12. **Sigal, A., Milo, R., Cohen, A., Geva-Zatorsky, N., Klein, Y., Alaluf, I., Swerdlin, N., Perzov, N., Danon, T., Liron, Y., Raveh, T., Carpenter, A. E., Lahav, G., and Alon, U.:** Dynamic proteomics in individual human cells uncovers widespread cell-cycle dependence of nuclear proteins, Nat. Methods, **3**, 525–531 (2006).
13. **Ansel, J., Bottin, H., Rodriguez-Beltran, C., Damon, C., Nagarajan, M., Fehrmann, S., Francois, J., and Yvert, G.:** Cell-to-cell stochastic variation in gene expression is a complex genetic trait, PLoS Genet., **4**, e1000049 (2008).
14. **Tseng, W. C., Haselton, F. R., and Giorgio, T. D.:** Transfection by cationic liposomes using simultaneous single cell measurements of plasmid delivery and transgene expression, J. Biol. Chem., **272**, 25641–25647 (1997).
15. **Cohen, R. N., van der Aa, M. A., Macaraeg, N., Lee, A. P., and Szoka Jr., F. C.:** Quantification of plasmid DNA copies in the nucleus after lipoplex and polyplex transfection, J. Control. Release, **135**, 166–174 (2009).
16. **Hakamada, K., Okamoto, M., and Hanai, T.:** Novel technique for preprocessing high dimensional time-course data from DNA microarray: mathematical model-based clustering, Bioinformatics, **22**, 843–848 (2006).
17. **Tomida, S., Hanai, T., Honda, H., and Kobayashi, T.:** Analysis of expression profile using fuzzy adaptive resonance theory, Bioinformatics, **18**, 1073–1083 (2002).
18. **Rousseeuw, P.:** Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, J. Comput. Appl. Math., **20**, 53–65 (1987).
19. **Chang, H. H., Hemberg, M., Barahona, M., Ingber, D. E., and Huang, S.:** Transcriptome-wide noise controls lineage choice in mammalian progenitor cells, Nature, **453**, 544–547 (2008).
20. **Margalit, A., Vlcek, S., Gruenbaum, Y., and Foisner, R.:** Breaking and making of the nuclear envelope, J. Cell. Biochem., **95**, 454–465 (2005).
21. **Escriou, V., Carriere, M., Bussone, F., Wils, P., and Scherman, D.:** Critical assessment of the nuclear import of plasmid during cationic lipid-mediated gene transfer, J. Gene Med., **3**, 179–187 (2001).
22. **Mortimer, I., Tam, P., MacLachlan, I., Graham, R. W., Saravolac, E. G., and Joshi, P. B.:** Cationic lipid-mediated transfection of cells in culture requires mitotic activity, Gene Ther., **6**, 403–411 (1999).
23. **Miller, A. M. and Dean, D. A.:** Cell-specific nuclear import of plasmid DNA in smooth muscle requires tissue-specific transcription factors and DNA sequences, Gene Ther., **15**, 1107–1115 (2008).
24. **Kishida, T., Asada, H., Kubo, K., Sato, Y. T., Shin-Ya, M., Imanishi, J., Yoshikawa, K., and Mazda, O.:** Pleiotrophic functions of Epstein–Barr virus nuclear antigen-1 (EBNA-1) and oriP differentially contribute to the efficiency of transfection/expression of exogenous gene in mammalian cells, J. Biotechnol., **133**, 201–207 (2008).
25. **Fischer, N., Kremmer, E., Lautscham, G., Mueller-Lantzsch, N., and Grasser, F. A.:** Epstein–Barr virus nuclear antigen 1 forms a complex with the nuclear transporter karyopherin alpha2, J. Biol. Chem., **272**, 3999–4005 (1997).
26. **Remington, S. J.:** Negotiating the speed bumps to fluorescence, Nat. Biotechnol., **20**, 28–29 (2002).
27. **Tsien, R. Y.:** The green fluorescent protein, Annu. Rev. Biochem., **67**, 509–544 (1998).
28. **Heim, R., Cubitt, A. B., and Tsien, R. Y.:** Improved green fluorescence, Nature, **373**, 663–664 (1995).

Published in final edited form as:
Nature. 2010 April 15; 464(7291): 993–998. doi:10.1038/nature08987.

International network of cancer genome projects

The International Cancer Genome Consortium*

Abstract

The International Cancer Genome Consortium (ICGC) was launched to coordinate large-scale cancer genome studies in tumors from 50 different cancer types and/or subtypes that are of clinical and societal importance across the globe. Systematic studies of over 25,000 cancer genomes at the genomic, epigenomic, and transcriptomic levels will reveal the repertoire of oncogenic mutations, uncover traces of the mutagenic influences, define clinically-relevant subtypes for prognosis and therapeutic management, and enable the development of new cancer therapies.

The genomes of all cancers accumulate somatic mutations[1]. These include nucleotide substitutions, small insertions and deletions, chromosomal rearrangements and copy number changes that can affect protein-coding or regulatory components of genes. In addition, cancer genomes usually acquire somatic epigenetic "marks" compared to non-neoplastic tissues from the same organ, notably changes in the methylation status of cytosines at CpG dinucleotides.

A subset of the somatic mutations in cancer cells confers oncogenic properties such as growth advantage, tissue invasion and metastasis, angiogenesis, and evasion of apoptosis[2]. These are termed "driver" mutations. The identification of driver mutations will provide insights into cancer biology and highlight novel drug targets and diagnostic tests. Knowledge of cancer mutations has already led to the development of specific therapies, such as trastuzumab for HER2/neu positive breast cancers[3] and imatinib, which targets BCR-ABL tyrosine kinase for the treatment of chronic myeloid leukemia[4,5]. The remaining somatic mutations in cancer genomes that do not contribute to cancer development are called "passengers". These mutations provide insights into the DNA damage and repair processes that have been operative during cancer development, including exogenous environmental exposures[6,7]. In most cancer genomes, it is anticipated that passenger mutations, as well as germline variants not yet catalogued in polymorphism databases, will substantially outnumber drivers.

Large-scale analyses of genes in tumors have revealed that the mutation load in cancer is abundant and heterogeneous[8-13]. Preliminary surveys of cancer genomes have already demonstrated their relevance in identifying new cancer genes that constitute potential therapeutic targets for several types of cancer, including PIK3CA[14], BRAF[15], NF1[10], KDR[10], PIK3R1[9], and histone methyltransferases and demethylases[16,17]. These projects have also yielded correlations between cancer mutations and prognosis, such as IDH1 and IDH2 mutations in several types of gliomas[13,18]. Advances in massively parallel sequencing technology have enabled sequencing of entire cancer genomes [19-22].

Following the launch of comprehensive cancer genome projects in the United Kingdom (Cancer Genome Project)23 and the United States (The Cancer Genome Atlas)24, cancer

Author Information: Correspondence should be addressed to T.J.H. (tom.hudson@oicr.on.ca).
*List of participants and affiliations appear at the end of the paper.

genome scientists and funding agencies met in Toronto (Canada) in October 2007 to discuss the opportunity to launch an international consortium. Key reasons for its formation were: (1) the scope is huge; (2) independent cancer genome initiatives could lead to duplication of effort or incomplete studies; (3) lack of standardization across studies could diminish the opportunities to merge and compare datasets; (4) the spectrum of many cancers is known to vary across the world; (5) an international consortium will accelerate the dissemination of datasets and analytical methods into the user community.

Working groups were created to develop strategies and policies that would form the basis for participation in the ICGC. The goals of the Consortium (Box 1) were released in April 2008 (http://www.icgc.org/files/ICGC_April_29_2008.pdf). Since then, working groups and initial member projects have further refined the policies and plans for international collaboration.

Bioethical Framework

ICGC members agreed to a core set of bioethical elements for consent as a precondition of membership (Box 2). The Ethics and Policy Committee has created patient consent templates for both prospective collection and retrospective use of samples and data for ICGC projects. Differences in project-specific requirements and national legal frameworks may require some local amendments, while still reflecting the core principles of ICGC.

The ICGC recognizes a delicate balance between protecting participants' personal data and sharing these data to accelerate cancer research. Data access policies have been drawn up that are respectful of the rights of the donors, while allowing ICGC data derived from samples to be shared ethically among a wide research community. Two levels of access have been implemented. For data that cannot be used to identify individuals, "Open access" datasets are publically available. These include data such as gender, age range, histology, normalized gene expression values, epigenetic datasets, somatic mutations, summaries of germline data, and study protocols. "Controlled access" datasets contain germline genomic data and detailed clinical information that are associated to a unique individual whose personal identifiers have been removed. To access controlled datasets researchers must seek authorizations by contacting the Data Access Compliance Office (DACO) (http://www.icgc.org/daco). An independent International Data Access Committee (IDAC) oversees the work of the DACO and provides assistance with resolving issues that arise.

Pathology and Clinical Annotation

Large-scale genomic studies of human tumors rely on the availability of fresh frozen tumor tissue. To address the paucity of samples that meet ICGC standards, many projects have initiated prospective collections of high quality source material. Accordingly, the ICGC recommended procedures to promote consistency of sample processing throughout the Consortium and ensure a series of quality features such as high tissue integrity and tumor cell content. Each project will need to include diverse data types such as environmental exposures, clinical history of participants, tumor histopathology, and clinical outcomes.

Tumors display considerable clinical and biological heterogeneity which has resulted in a variety of tumor classifications. Within the ICGC, special measures are taken to promote the consistency of diagnosis. These include the coordination of diagnostic criteria among groups investigating tumors that are related, and policies that all samples will be reviewed by at least two independent reference pathologists. Furthermore, images of the stained tumor sections (or blood smear or cytospins for hematological neoplasias), from which diagnoses were made, will be stored and made available to the community.

Although different tumor types may require specific procedures for tumor acquisition or compilation of clinical and environmental data, ICGC has set guidelines regarding the use of common definitions and data standards. This will allow ICGC data users to identify correlations between tumor-specific molecular changes with clinical and histopathological data including prognosis, prediction of therapy response and tumor classification schemes for diagnosis.

Study Design and Statistical Issues

To identify cancer-related genes, one needs to detect genes that are mutated at a higher frequency than the background mutation rate. Given that several driver genes have been found to be mutated at low frequencies, ICGC will identify somatic mutation observed in at least 3% of tumors of a given subtype. ICGC determined that 500 samples would be needed per tumor type (although for rare tumor types, a smaller sample size may be justified). In practice, the degree of heterogeneity of a given tumor type is difficult to know in advance, such that some particularly heterogeneous tumor types may require larger sample collections.

Cancer Genome Analyses

High-quality catalogues of somatic mutations from whole cancer genomes will ultimately be the ICGC standard. Shotgun sequencing employing second generation technologies can detect all classes of somatic mutation implicated in cancer. Moreover, if the level of coverage is sufficient, comprehensive high quality catalogues of somatic mutations from individual cancer genomes can be acquired with >90% sensitivity and >95% specificity. In order to achieve this, it will be necessary to sequence both the genome of the cancer and of a normal tissue from the same individual to distinguish germline variants. Although a few genomes of this standard have already been generated, the cost and the continuing technology development will mean that interim analyses of particularly informative sectors of the genome will be carried out, for example of all coding exons and microRNAs.

For each individual cancer genome, the catalogue of somatic mutations will be supplemented by genome-wide information on the state of methylation of CpG dinucleotides. The optimal strategies and technologies to achieve this are not yet clear. Moreover, the genomes of individual cancers will be accompanied, where possible, by analyses of the transcriptome. Although conventional array-based approaches currently predominate, it is preferable that RNA sequencing becomes the standard as sequencing has a greater dynamic range[25] and provides additional information including novel transcripts and sequence variants[26].

ICGC Datasets

The distributed nature of the Consortium coupled with the large size of the datasets makes it cumbersome to store all data in a single centralized repository. For this reason, the ICGC has adopted a "franchise" database model for integrating the information and making it available to the public. Under this model, each member project releases tumor information by copying it into its local franchise database after it has been quality checked. Each franchise database shares a common schema to describe the specimens, the associated clinical information, and their genome characterization data. ICGC primary data files, including sequencing traces, are sent to the National Center for Biotechnology Information (NCBI) and/or the European Bioinformatics Institute (EBI) for archiving, while interpreted data sets, such as somatic mutation calls, are stored in franchise databases. The ICGC franchise databases and web portal use BioMart27, a data federation technology originally developed for use in Ensembl28, and since adopted for use by multiple model organism and genome databases. The management of the ICGC data flow is the responsibility of the ICGC Data Coordination Center (DCC) located at the Ontario Institute for Cancer Research.