ORIGINAL ARTICLE

Establishment and management of a lung cancer biobank in Eastern China

Keke Yu[1,2], Jie Zhang[2], Xin Li[1], Lei Xu[1], Ye Zhang[1,2], Jie Xing[1,2] Jinchen Shao[2], Lei Zhu[2], Jinguo Liu[2], Lanxiang Zhao[2] & Baohui Han[1,3]

1 Department of Biaobank, Shanghai Chest Hospital, Shanghai JiaoTong University, Shanghai, China
2 Department of Pathology, Shanghai Chest Hospital, Shanghai JiaoTong University, Shanghai, China
3 Department of Pulmonary Medcine, Shanghai Chest Hospital, Shanghai JiaoTong University, Shanghai, China

Abstract

Background: The prevalence of lung cancer, a highly complex neoplasm, increases annually. Thus, a lung cancer biobank, which stores lung cancer tissue and blood matched according to standard methods, is needed to advance lung cancer research and develop promising therapies.

Methods: To accomplish this aim, we implemented standardized procedures for tissue samples and patient information acquired from consenting donors. The banked tissue includes blood, pleural effusions, and surgical resection samples. An independent information management system was used to match samples and collect data, including clinical cancer manifestation, laboratory tests, and de-identified data about cancer patients.

Results: From 2009 to 2013, more than 2000 lung cancer cases were collected. At this time, we have more than 10 000 biological samples stored in our biobank. DNA, ribonucleic acid (RNA), and protein quality were confirmed to be appropriate for clinical and basic research.

Conclusion: Our standardized, large-scale lung cancer biobank offers high quality cancer research samples for China and the world.

Introduction

Biobanks are increasingly needed to study human cancer to facilitate basic and clinical research and to develop novel ways to achieve early diagnostics, prevention, and personalized treatment. [1] Lung cancer is the most common cancer in the world, representing almost 12% of all new cancer cases.[2,3] The incidence of lung cancer in Shanghai has increased over the past decade and contributes to 26% of female and 31% of male cancer-related deaths. As such, it is the greatest cause of cancer-related mortality.[4,5] Air pollution may contribute to lung cancer etiology; therefore, the incidence of lung cancer in China is likely to rise.[6,7] Research into the mechanisms of lung cancer development is an area of special concern, and advances in cancer treatment programs depend on high quality tissue samples. A lung cancer biobank in Eastern China will facilitate understanding of cancer etiology and assist investigators in elucidating features of cancer that can be exploited to create new therapeutic targets.

Increasingly, new chemotherapeutic agents depend on demonstrated efficacy in human tumor tissue or cell cultures.[8] Different responses to lung cancer treatment are often attributable to racial and genetic differences. In fact, the frequency of lung cancer histological subtypes has changed over the past few decades. The Shanghai Chest Hospital, affiliated with Shanghai JiaoTong University is a specialized tertiary level hospital, performing more than 4000 lung cancer surgical resection cases annually. Thus, this area has a wealth of lung cancer resources. Since 2009, we have formed a standardized process to collect biological samples and related clinical, pathological and follow-up information. Presently, the biobank at the Shanghai Chest Hospital is supported and funded by the Shanghai JiaoTong University "985 Project," (which is a biological sample library) and the Shanghai Science and Technology Commission. Ideally, these samples and related data will guide the selection of clinical treatment methods and can be used to estimate the prognosis of lung cancer patients.

Materials and methods

Tissue origins

The Shanghai Chest Hospital Ethics Committee approved the collection and informed consent protocols for the biological samples. Tissue collected in the lung cancer biobank is mainly from the Shanghai Chest Hospital, from patients with lung cancer who received surgical treatment. Samples from inflammatory pseudotumor, pulmonary fibrosis, and clear cell cancer were also collected as controls of benign diseases. Collection began in 2009 and included frozen lung cancer tissue (including normal matched tissues), formalin-fixed paraffin-embedded tissue, and blood specimen samples.

Blood sample collection

Blood samples were collected from donors who underwent surgery in our hospital. Blood was collected from each donor before surgery and was then centrifuged and the separated plasma/serum and buffy coat were collected and labeled with a barcode before being stored at −80°C.

Frozen tissue specimen collection

Surgical specimens were at once snap-frozen in liquid nitrogen *in situ* after resection to minimize hypoxic phenomena on genetic expression.[9] Tissues were initially identified by a pathologist as malignant and then decisions were made about separating tissue fragments for diagnosis. Tissues with necrotic areas and/or massive ischemia were not collected and aseptic conditions were used. Also, normal matched tissue from the same organ (adjacent tissue: 2–5 cm to the cancer margin and distant tissue more than 5 cm near the cancer margin) was obtained, frozen, and fixed for paraffin embedding. Samples were cut it into eight equal portions, each of which was not less than 0.5 cm.[3] Samples were then frozen in liquid nitrogen. Components of such specimens can be used for DNA, ribonucleic acid (RNA) and protein research.[10,11]

Nucleic acid extraction and identification

We selected frozen, formalin-fixed, paraffin-embedded tissue treated with RNAlater and blood samples stored at −80°C. Genomic DNA and total RNA were extracted using nucleic acid extraction kits. Absorbances and concentrations were measured with a NanoDrop 1000 UV spectrophotometer. Data integrity was confirmed using 1% agarose gel electrophoresis and an Agilent Bioanalyzer 2100.

Protein extraction and identification

Cell lysate and phenylmethylsulfonyl fluoride (PMSF) were added to 100 mg frozen tissue samples and these were placed on ice for 30 minutes. Samples were centrifuged for 30 minutes at 4°C at 12 000 rpm, and supernatant was collected. Protein quality was confirmed by Western blot.

Information management

The "985"biobank project of Shanghai JiaoTong University has unified coding rules for donors and samples. These codes are implemented with a two-dimensional code label that contains basic sample information to ensure orderly sample management. Databases house personal patient information such as name, gender, age, clinical record number, pathology reference number, sample number, sample amount, and storage location. The samples were re-examined daily to ensure quality.

Sample pathology and clinical follow-up information integration

Lung cancer has many pathological types and gene mutations that contribute to the cancer. Thus, there are different clinical treatments and potential prognoses. To keep this information reliably catalogued, electronic medical records from the hospital are used.

Database management

The database is interfaced with specific software for the biobank, and this software is run daily to integrate any new information within the biobank. Frozen tissue specimens, formalin-fixed paraffin-embedded tissues, blood specimens, clinically relevant material, and long-term follow-up data comprise the first grade data. Serial detection of DNA, RNA and protein from the first grade data make up the second grade data. Third-grade data consist of pathological information, such as molecular mutations (epidermal growth factor receptor [EGFR] and K-ras mutation), and clinically long-term follow-up data. These data guide clinical treatment method selection and can be helpful for judging the prognosis of lung cancer patients.

One person presides over the biobank and samples are requested by a formal application process. Sample requests are verified and submitted to the Hospital Biobank Management Committee for approval. After experiments using biobank material are complete, investigators return remaining samples and any information associated with them. All samples are de-identified of all patient information to protect patient privacy.

Results

Biobank specimen constituents

Since 2009, 2007 cancer cases and 16 546 samples have been banked (6016 frozen tissues, 2000 tissue blocks, and 8530
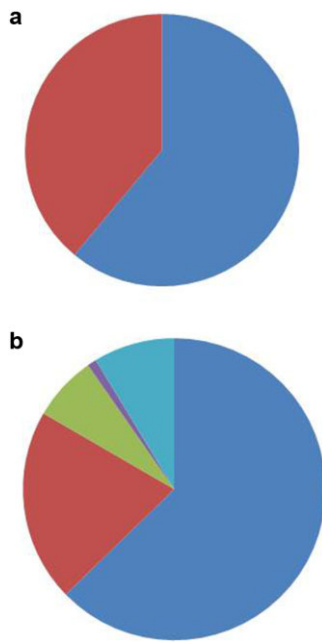
Figure 1 Specimen constituents. (**a**) Constituent ratio of gender was 61:39 (male vs. female patients); (**b**) Lung adenocarcinoma and lung squamous cell carcinoma accounted for the majority of the total surgical specimens: adenocarcinoma specimens represented 62.72% and squamous cell carcinomas comprised 20.66% of the samples. Small cell lung cancer patients may not be eligible for surgery because of early transfer and is often treated by radiotherapy and chemotherapy. (**a**) ■, male 61.08%; ■, female 38.92% (**b**) ■, adenocarcinoma 62.72%; ■, squamous cell carcinoma 20.66%; ■, small cell carcinoma 6.96%; ■, large cell carcinoma 0.97%; ■, other types 8.69%.

blood specimens). Of the cancer cases, 61.08% were samples from males.

Tumor types

Lung adenocarcinoma and lung squamous cell carcinoma accounted for the majority of all surgical specimens: adenocarcinoma specimens represented 62.72% and squamous cell carcinoma comprised 20.66% of the samples. Small cell lung cancer represented a smaller proportion (6.96%) of surgical specimens, and blood specimens were chiefly collected for this type of cancer. Large cell carcinoma specimens represented 0.97% of surgical resection specimens. Other samples included adenosquamous carcinoma, sarcoma, adenoid cystic carcinoma, and carcinosarcoma; cumulatively, they represented 8.69% of the surgical resection specimens. (Fig. 1)

Hematoxylin and eosin (H&E) staining

Twenty cases were randomly selected from lung adenocarcinoma, squamous cell carcinoma and normal control tissues

embedded in optimum cutting temperature compound (OCT). These were stained with hematoxylin and eosin and observed under a microscope. Tumor cells comprised more than 80% of the specimens. (Fig. 2)

DNA quality

Ten tumor tissue samples were randomly selected and genomic DNA was extracted and evaluated with a NanoDrop 1000 UV spectrophotometer. A260/280 values were ~2.0 (Table 1). Bands were clearly visible after 1% agarose gel electrophoresis (Fig. 3) and DNA was well preserved.

Ribonucleic acid(RNA) quality

Ten tumor tissue samples were randomly selected and total RNA was extracted and evaluated with a NanoDrop 1000 UV spectrophotometer and Agilent Bioanalyzer 2100. A260/280 values were ~2.0 (Table 1). Bands were clearly visible after 1% agarose gel electrophoresis (Fig. 4). The RNA Integrity Number (RIN) was greater than seven. RNA was well preserved and could be used in subsequent scientific experiments.

Protein quality

Ten tumor tissue samples were randomly selected from liquid nitrogen preservation and protein was extracted. Rabbit anti-human actin antibody was the primary antibody (1:1000 dilution), and the target band was detected by Western blot (Fig. 5); the target band was clearly visible.

Discussion

Research biobanks are organized to collect and store biological samples for scientific research. Malignant tumors are significant public health problems, and research indicates that improving early diagnosis and rational tumor treatment are keys to prolonging patient survival and improving quality of life.[12,13] As important resources for translational and personalized medicine, tumor biological samples are valuable and non-renewable resources, which include actual samples with phenotypic and biological information. Rich phenotypic and biological information directly determine the value of any biobank.[14,15] Integrity and high quality of tumor tissue sample preservation systems matter, and each sample should be regarded as a small genetic information platform, to be used not only as a resource for existing research programs, but also for serving as an information resource database for future basic and clinical cancer research. Formerly known as the lung cancer tissue bank, since 2009, the Shanghai Chest Hospital biobank has collected and preserved tumor tissues, normal matched tissues, corresponding blood, pleural effu-
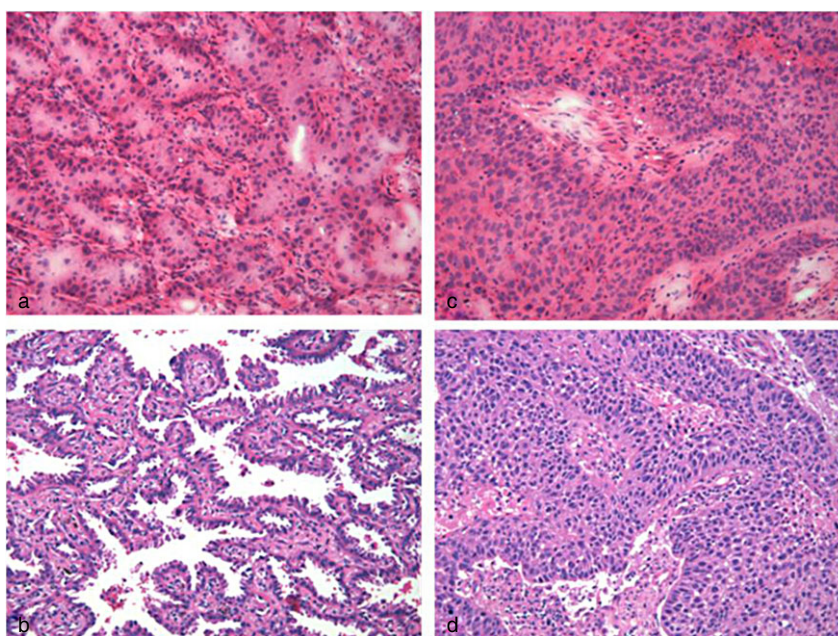
Figure 2 Hematoxylin and eosin (H&E) staining. (**a**) Frozen adenocarcinoma samples stained with H&E; (**b**) Formalin fixed paraffin embedded (FFPE) sample from the same donor stained with H&E. (**c**) Frozen squamous cell carcinoma sample stained with H&E; (**d**) FFPE sample from the same donor stained with H&E. (×200).

sion, and other biological samples, as well as complete clinical histories of diagnosis, treatment, and patient follow-up data in a manner that ensured quality control. These samples have been used for scientific research and academic exchanges and 82 clinical studies have been undertaken with biobank information in the past five years.

Tissue and blood in the lung cancer biobank were often used for molecular biology experiments, such as polymerase chain reaction, genechip, epigenetic analysis and gene mutation studies. Therefore, the proportion of tumor cells in each sample is important. A minimum threshold of 65% tumor cells is required for molecular biology investigations. In this study, the pathologist stained samples with hematoxylin and eosin to estimate sampling accuracy. A pathologist confirmed that tumor cells comprised more than 70% of the collected samples in a biobank quality control project.

DNA and RNA quality is directly related to follow-up experiment reliability and repeatability.[16] In addition, RNA in collected tissue is easily degraded.[17] Thus, RNA detection is an important indicator of quality control. Key criteria for RNA identification are integrity and uniformity, and the

Table 1 OD values and concentration of biological specimen DNA and RNA from different storage methods

DNA	OD260/280	Ng/ul	RNA	OD260/280	Ng/ul
1	1.93	106.48	1	1.91	132.65
2	1.95	144.33	2	1.95	131.9
3	1.87	81.56	3	1.98	199.43
4	1.89	98.92	4	1.97	187.43
5	1.97	147.56	5	1.95	140.15
6	1.81	90.97	6	1.98	178.45
7	1.93	98.08	7	1.91	150.71
8	1.8	81.56	8	1.92	133.73
9	1.87	58.34	9	1.89	111.63
10	1.83	86.56	10	1.93	128.47
11	1.89	82.15	11	1.82	79.43
12	1.81	75.27	12	1.83	85.71
13	1.91	80.99	13	1.84	88.64
14	1.81	74.82	14	1.81	74.01
15	1.86	76.49	15	1.82	80.48

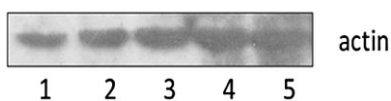OD, optical density; RNA, ribonucleic acid.
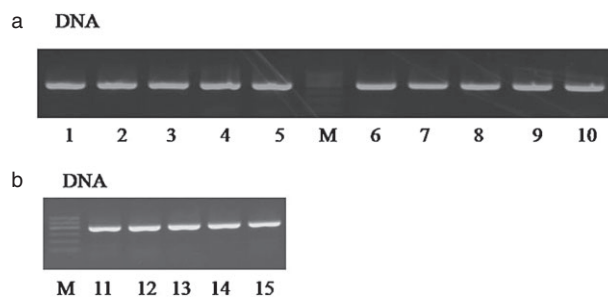
Figure 3 DNA quality. (**a**) 1% agarose gel electrophoresis for genomic DNA;1–5 were extracted from frozen tissue samples, 6–10 were from the same donor's FFPE tissue samples. (**b**) 11–15 were extracted from the same donor's frozen blood samples.
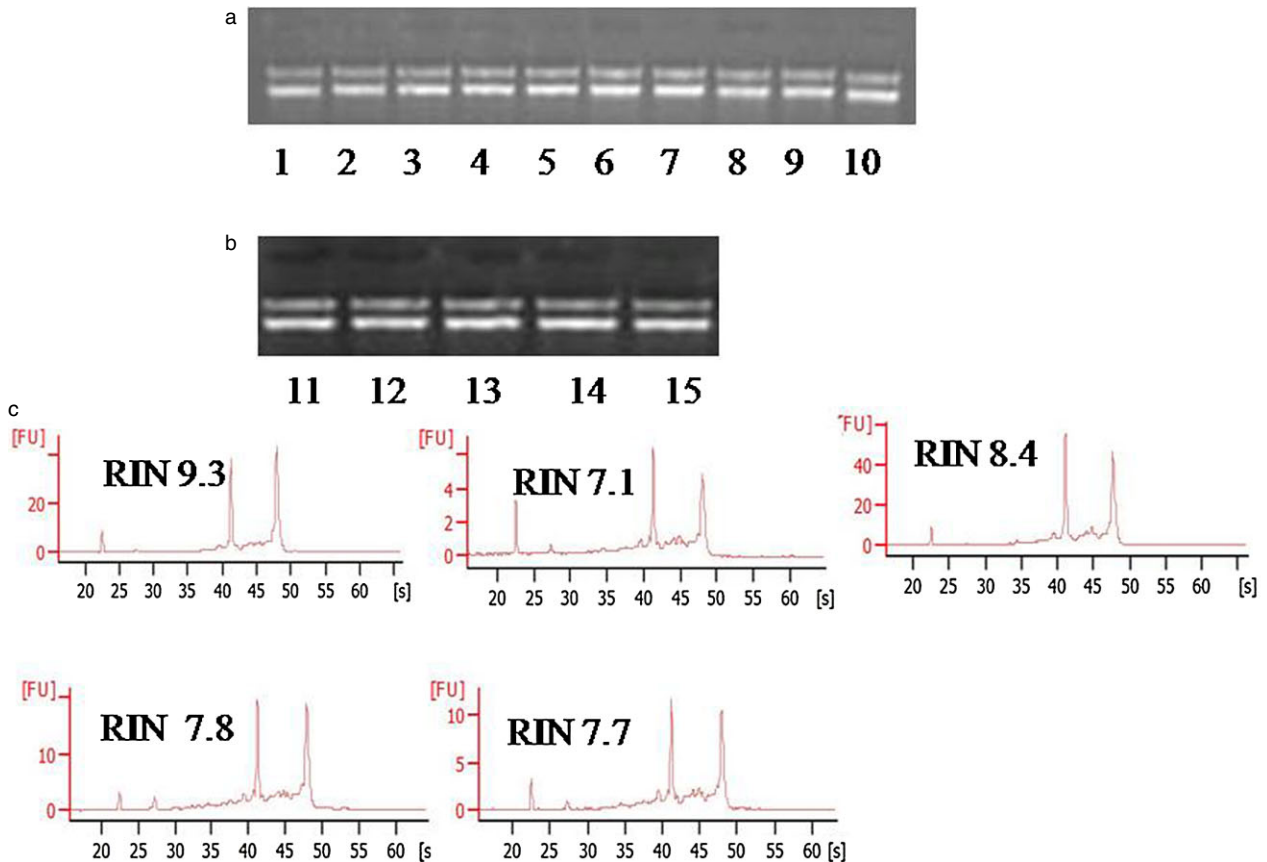
Figure 4 Ribonucleic acid (RNA) quality. (**a**) 1% agarose gel electrophoresisfor total RNA; 1–5 were extracted from frozen tissue samples, 6–10 were from the same donor's FFPE tissue samples. (**b**) 11–15 were extracted from the same donor's frozen blood samples. (**c**) TheRNA Integrity Number (RIN) evaluated with a NanoDrop 1000 ultraviolet spectrophotometer was greater than seven. See Figure 3 for the order of tested samples.

most economic and effective method for confirming this is agarose gel electrophoresis and RIN.[18,19] To ensure biological sample quality and prevent DNA and RNA degradation, tissue samples were preserved and stored within 30 minutes of collection and blood samples were stored within two hours of collection. RNAlater was used to preserve tumor tissue to prevent degradation and to avoid repeated freezing and thawing. We dispensed the biological samples by national standards and increased the numbers of samples stored. We confirmed the quality of DNA, RNA and protein at the end of the sampling month. In addition, the Shanghai Science and Technology Commission annually performed quality control studies on the biobank samples. Quality testing ensured that RNA, DNA, and protein in the biobank

samples were well preserved and could be used for scientific research. Specifically, our biobank samples, along with relevant matched electronic medical records, permit clinical, pathological, molecular pathology, treatment, long term follow-up and laboratory test data to be used for studies. Meanwhile, we intend to construct a lung cancer information management system in our hospital. Because patient follow up over long periods is very difficult, we initiated the management system in the department of pulmonary medicine first. We intend to expand the system to cover the entire hospital so that we can develop a comprehensive follow-up system that improves electronic medical records and maintains biobank database reliability. The Shanghai JiaoTong University regularly provides maintenance and upgrade services for the biobank database. These samples and their associated information can facilitate tumor-targeted research and allow exploration into treatment resistance, because treatment data, including repeated chemotherapy, radiotherapy, and follow-up information, are also included. Samples can be easily located with simple word queries and this database is available to hospital staff, so it offers

Figure 5 Protein quality. From frozen samples, 1–5 protein was extracted and tested via Westernblot.

convenience for researchers and collaborators to conduct scientific research and further analyses.

The goal for the lung cancer biobank was to facilitate biomedical research using tissues collected from Eastern Chinese patient populations. Because Eastern and Western populations have different responses to lung cancer treatment, our biobank can expand studies into genetic differences in treatment responses. Banked DNA, RNA, protein, serum/plasma specimens and tissue microarrays will enable these types of investigations.

Conclusion

Our report offers a template for other biobanks, suggesting methods for selecting varied samples and managing such histological subtypes with the optimal quality. The Biobank of the Shanghai Chest Hospital is supported and funded by the Shanghai JiaoTong University "985 Project" biological sample library and the Shanghai R&D public service platform project "common malignant tumor biobank (Network) construction." At this time, we can share 2000 cases of biological lung cancer samples, allowing us to not only meet the needs of our research, but also to provide samples for other scientists.

Acknowledgments

Disclosure

No authors report any conflict of interest.

References

1 Li HL, Qiu Y, Zhang X *et al.* Establishment of a lung cancer biobank of a southern Chinese population. *J Thorac Dis* 2009; **1**: 17–22.

2 Zhuo YL, Guo QS, Song PP *et al.* Correlation study and significance of the EGFR expression in serum, lymph nodes and tumor tissue of NSCLC. *Thorac Cancer* 2014; **5**: 31–7.

3 Youlden DR, Cramb SM, Baade PD. The international epidemiology of lung cancer: geographical distribution and secular trends. *Thorac Oncol* 2008; **3**: 819–31.

4 Zhang Y, Jin B, Shao M *et al.* Monitoring of carcinoembryonic antigen levels is predictive of EGFR mutations and efficacy of EGFR-TKI in patients with lung adenocarcinoma. *Tumour Biol* 2014; **5**: 4921–8.

5 Jemal A, Siegel R, Ward E, Murray T, Xu J, Thun MJ. Cancer statistics, 2007. *CA Cancer J Clin* 2007; **57**: 43–66.

6 Pulley JM, Brace MM, Bernard GR, Masys DR. Attitudes and perceptions of patients towards methods of establishing a DNA biobank. *Cell Tissue Bank* 2008; **9**: 55–65.

7 Wang QS, Lin XP. Lung cancer epidemiology in mainland China. *Chin J Clin Oncol* 2006; **3**: 244–53.

8 Vaught JB, Caboux E, Hainaut P. International efforts to develop biospecimen best practices. *Cancer Epidemiol Biomarkers Prev* 2010; **19**: 912–15.

9 Bao WG, Zhang X, Zhang JG *et al.* Biobanking of fresh-frozen human colon tissues: impact of tissue ex-vivo ischemia times and storage periods on RNA quality. *Ann SurgOncol* 2013; **20**: 1737–44.

10 Yuille M, Dixon K, Platt A *et al.* The UK DNA banking network: a "fair access" biobank. *Cell Tissue Bank* 2010; **11**: 241–51.

11 Ruíz-Godoy L, Meneses-García A, Suárez-Roa L *et al.* Organization of a tumor bank: the experience of the National Cancer Institute of Mexico. *Pathobiology* 2010; **77**: 147–54.

12 Nather A, Phillips GO, Morales Pedraza J, Lee CC. Training tissue bank operators: the International Atomic Energy Agency (IAEA)/National University of Singapore (NUS) 10 years of experience. *Cell Tissue Bank* 2009; **10**: 183–90.

13 Van Meerbeeck JP, Fennell DA, Ruysscher DK. Small-cell lung cancer. *Lancet* 2011; **378**: 1741–55.

14 Grizzle W, Bell WC, Sexton KC. Issues in collecting, processing and storing human tissues and associated information to support biomedical research. *Cancer Biomark* 2010; **9**: 531–49.

15 Johnsen IK, Hahner S, Brière JJ *et al.* Evaluation of a standardized protocol for processing adrenal tumor samples: preparation for a European adrenal tumor bank. *Horm Metab Res* 2010; **42**: 93–101.

16 Hall GL, Kademani D, Risk JM, Shaw RJ. Tissue banking in head and neck cancer. *Oral Oncol* 2008; **44**: 109–15.

17 Andreasson A, Kiss NB, Juhlin CC, Höög A. Long-term storage of endocrine tissue at −80°C does not adversely affect RNA quality or overall histomorphology. *Biopresery Biobank* 2013; **11**: 366–70.

18 Ribeiro Olivieri EH, de Andrade Franco L, Pereira RG, Carvalho Mota LD, Campos AH, Carraro DM. Biobankingpractice: RNA storage at low concentration affects integrity. *Biopresery Biobank* 2014; **12**: 46–52.

19 Schroeder A, Mueller O, Stocker S *et al.* The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol* 2006; **7**: 3.