

Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases

Xi Chen^{1,*}, Yi Ba^{2,*}, Lijia Ma^{3,4,*}, Xing Cai¹, Yuan Yin¹, Kehui Wang¹, Jigang Guo¹, Yujing Zhang¹, Jiangning Chen¹, Xing Guo¹, Qibin Li^{3,4}, Xiaoying Li⁵, Wenjing Wang⁶, Yan Zhang¹, Jin Wang¹, Xueyuan Jiang¹, Yang Xiang¹, Chen Xu¹, Pingping Zheng³, Juanbin Zhang³, Ruiqiang Li³, Hongjie Zhang¹, Xiaobin Shang², Ting Gong², Guang Ning⁵, Jun Wang^{3,4}, Ke Zen¹, Junfeng Zhang¹, Chen-Yu Zhang¹

¹Jiangsu Diabetes Center, State Key Laboratory of Pharmaceutical Biotechnology, School of Life Sciences, Nanjing University, 22 Hankou Road, Nanjing, Jiangsu 210093, China; ²Tianjin Medical University Cancer Institute and Hospital, Tianjin 300000, China; ³Beijing Genomics Institute, Shenzhen 518083, China; ⁴Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 101300, China; ⁵Department of Endocrine and Metabolic Diseases, Ruijin Hospital Affiliated to The Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China; ⁶Shanghai Municipal Center for Disease Control and Prevention, Shanghai 200336, China

Dysregulated expression of microRNAs (miRNAs) in various tissues has been associated with a variety of diseases, including cancers. Here we demonstrate that miRNAs are present in the serum and plasma of humans and other animals such as mice, rats, bovine fetuses, calves, and horses. The levels of miRNAs in serum are stable, reproducible, and consistent among individuals of the same species. Employing Solexa, we sequenced all serum miRNAs of healthy Chinese subjects and found over 100 and 91 serum miRNAs in male and female subjects, respectively. We also identified specific expression patterns of serum miRNAs for lung cancer, colorectal cancer, and diabetes, providing evidence that serum miRNAs contain fingerprints for various diseases. Two non-small cell lung cancer-specific serum miRNAs obtained by Solexa were further validated in an independent trial of 75 healthy donors and 152 cancer patients, using quantitative reverse transcription polymerase chain reaction assays. Through these analyses, we conclude that serum miRNAs can serve as potential biomarkers for the detection of various cancers and other diseases.

Keywords: serum-microRNA, expression profile, fingerprint, blood-based biomarker, Solexa, cancers, diabetes

Cell Research advance online publication 2 September 2008; doi: 10.1038/cr.2008.282

Introduction

Cancers, including lung cancer and colorectal cancer, are often diagnosed at a late stage with concomitant poor prognosis [1-4]. Although tumor markers greatly improve diagnosis, the invasive, unpleasant, and inconvenient nature of current diagnostic procedures limits their application [3, 4]. Hence, there is a great need for identification

of novel non-invasive biomarkers for early tumor detection. miRNAs, a class of naturally occurring small non-coding RNAs of 19-25 nucleotides (19-25 nt) in length, have recently been linked to cancer development [5, 6]. Recently, altered miRNA expression has been reported in various cancers, and the profiles of tissue miRNAs exhibit great potential for an application in cancer definition [5, 6]. In the present study, we report the surprising and exciting discovery that serum and plasma contain a large amount of stable miRNAs derived from various tissues/organs, and that the expression profile of these miRNAs shows great promise as a novel non-invasive biomarker for diagnosis of cancer and other diseases.

Results

miRNAs are present in serum and plasma

*These three authors contributed equally to this work.

Correspondence: Chen-Yu Zhang^a, Junfeng Zhang^b, Ke Zen^c, Jun Wang^d
Tel/fax: +86-25-83686234

^aE-mail: cyzhang@nju.edu.cn

^bE-mail: jfzhang@nju.edu.cn

^cE-mail: kzen@nju.edu.cn

^dE-mail: wangj@genomics.org.cn

Received 30 July 2008; revised 7 August 2008; accepted 7 August 2008

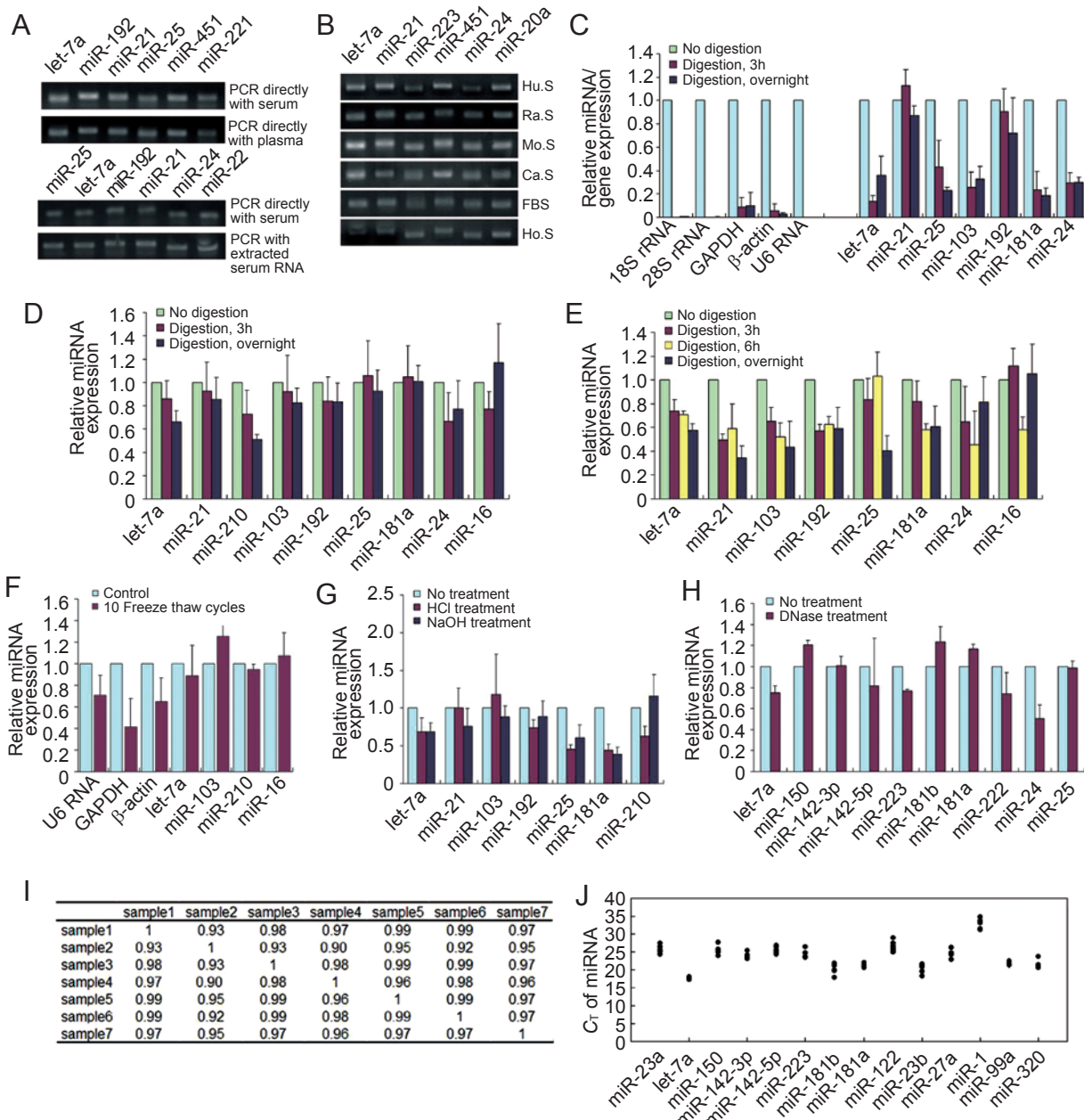


Figure 1 General characterization of miRNAs in serum and plasma. **(A)** The expression levels of indicated miRNAs evaluated by semi-quantitative RT-PCR analysis with 30 cycles. Upper panel: Serum and plasma of healthy subjects were directly used for semi-quantitative RT-PCR analysis without RNA extraction procedure. Lower panel: Comparison of serum miRNA expression by using semi-quantitative RT-PCR analysis with/without RNA extraction. **(B)** Semi-quantitative RT-PCR analysis of the indicated miRNAs in the serum from human (Hu.S), rat (Ra.S), mouse (Mo.S), calf (Ca.S), fetal bovine (FBS), and horse (Ho.S). **(C-G)** The stability of serum miRNAs. Total RNA was extracted from A549 cells and was then digested with RNase A for 3 h or overnight at 37 °C. The levels of miRNA or other RNAs were determined using qRT-PCR with 40 cycles (C). Extracted serum RNAs (D) and serum without RNA extraction (E) were treated with RNase A and showed no miRNA degradation. Serum was also subjected to 10 freeze-thaw cycles (F), or treated for 3 h in low (pH=1) or high (pH=13) pH solution (G). Then miRNA levels were evaluated by qRT-PCR with 40 cycles. **(H)** Serum was incubated with DNase I for 3 h and RNAs were analyzed by qRT-PCR with 40 cycles. **(I and J)** Comparison of levels of 14 miRNAs in the serum of 7 healthy subjects by qRT-PCR with 40 cycles. Pearson correlation coefficient indicated that the abundance of miRNAs in serum from different individual was quite correlated (I). The C_T values of each miRNA are illustrated (J).

A stem-loop reverse transcription polymerase chain reaction (RT-PCR) assay [7, 8] was adapted to screen mature miRNA expression in serum. As shown in the upper panel of Figure 1A, randomly selected miRNAs let-7a, miR-192, miR-21, miR-25, miR-451, and miR-221 were clearly expressed in both normal human serum and plasma as detected by semi-quantitative RT-PCR analysis with 30 cycles. In these experiments for detection of serum miRNAs, semi-quantitative RT-PCR analysis was performed directly using 10 μ l of sera without RNA extraction. As expected, almost identical results were obtained when semi-quantitative RT-PCR analysis was performed using extracted RNA from serum (Figure 1A, lower panel). The accuracy of semi-quantitative RT-PCR results was further validated by using traditional Sanger-based technology to sequence the PCR products of serum miRNAs. Out of 100 colonies obtained from PCR products of serum miRNAs, 87 showed the correct sequence of targeting miRNAs (data not shown). We also showed that miRNA let-7a, miR-21, miR-223, miR-451, miR-24, and miR-20a were detected in the serum of rats (Ra.S), mice (Mo.S), calves (Ca.S), bovine fetuses (FBS), and horses (Ho.S) (Figure 1B).

Serum miRNAs are resistant to RNase A digestion

Given the fact that serum contains ribonuclease, there should be no intact RNA in the serum. Presence of miRNAs suggests that these miRNA species are resistant to RNase digestion. Alternatively, the qRT-PCR products could be from contamination by degraded products of large molecular weight RNA, tRNAs, or genomic DNA. To assess these possibilities, vigorous studies were performed to characterize the stability of serum miRNAs. Firstly, some randomly selected miRNAs were amplified by RT-PCR from lung carcinoma A549 cell RNA extracts, with or without RNase A digestion. A group of large molecular weight RNAs, including 18s rRNA, 28s rRNA, GAPDH, β -actin, and U6, were used as controls. Surprisingly, miRNAs showed considerable resistance to the enzymatic cleavage by RNase A. For the miRNAs we tested, more than half of the molecules remained intact after 3 h of exposure to RNase A (Figure 1C). In contrast, all large molecular weight RNAs were rapidly degraded by RNase A, as was expected. Furthermore, serum or extracted serum RNAs were treated directly with RNase A. Strikingly, RNase A had hardly any effect on serum miRNAs with or without the RNA extraction procedure (Figure 1D and 1E). These data clearly demonstrate that miRNAs, particularly serum miRNAs, are resistant to RNase digestion. The stability of serum miRNAs was further studied by using sera, from various sources, after treatment under harsh conditions including boil-

ing, low/high pH, extended storage, freeze-thaw cycles, and so on. The qRT-PCR analysis of miRNAs in serum samples treated under these harsh conditions yielded no significant differences compared to non-treated serum. As shown in Figure 1F, serum miRNAs are resistant to multiple freeze-thaw cycles compared to large molecular weight RNAs. Moreover, serum miRNAs remained stable when serum was treated for 3 h in low (pH=1) or high (pH=13) pH solution (Figure 1G). To rule out the possibility of contamination by DNA fragments, qRT-PCR was performed in serum with or without DNase I treatment. Results showed that DNase I did not affect the level of serum miRNAs detected by qRT-PCR (Figure 1H). Moreover, in let-7a RT reaction, omitting reverse transcriptase resulted in no PCR product (data not shown). These results indicate that the qRT-PCR products are not a result of genomic DNA contamination.

Expression levels of serum miRNAs are reproducible and consistent among individuals

The variation of serum miRNAs among healthy human subjects was also characterized. Seven young (22-25 years) healthy Chinese subjects, four male and three female, were randomly selected and the expression levels of 14 miRNAs in their sera were analyzed by qRT-PCR. As shown in Figure 1I and 1J, expression levels of serum miRNAs among healthy subjects were quite consistent, with a Pearson correlation coefficient (R) close to 1. Taken together, these results show that serum miRNAs are stable and that their levels are reproducibly consistent among individuals of the same species.

miRNAs are the major fraction of small nucleotide species in serum as determined by Solexa sequencing

To exclude the possibility that the qRT-PCR miRNA products resulted from other small RNAs or degraded RNA fragments, and to further investigate the serum miRNA expression profile in normal human subjects, Solexa was employed to sequence all small RNAs (<30 nt) found in the serum of healthy subjects. Solexa is a high-throughput sequencing technology producing highly accurate, reproducible, and quantitative readouts of small RNAs, including those expressed at low levels [9]. Small serum RNAs from patients were sequenced simultaneously (the details are discussed in Figure 3). Direct sequencing by Solexa can unambiguously distinguish miRNAs from other small RNAs, such as tRNAs or degraded RNA fragments. For Solexa sequencing, the sera from 11 male and 10 female healthy Chinese subjects were pooled, separately, and total RNA of two pooled sera was extracted. Figure 2A shows the length distribution of small RNA reads in the pooled sera of healthy

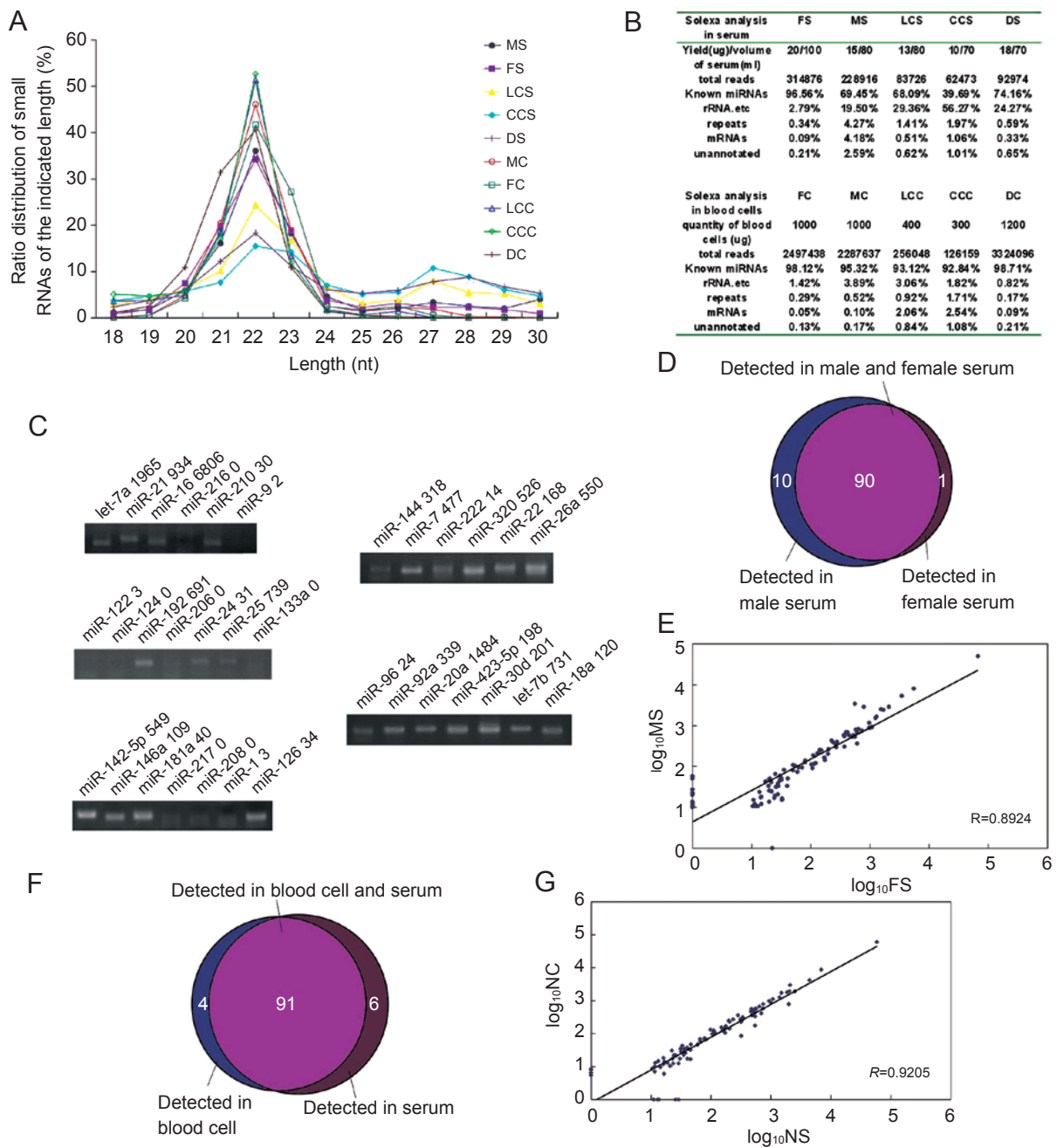


Figure 2 Solexa sequencing analysis of serum miRNAs in healthy subjects. **(A)** The distribution of small RNAs of various lengths (18-30 bp) sequenced by Solexa. Small RNAs were isolated from the serum of healthy male subjects (MS), female subjects (FS), non-small cell lung cancer patients (LCS), colorectal cancer patients (CCS), and diabetes patients (DS), as well as blood cells of healthy male subjects (MC), female subjects (FC), non-small cell lung cancer patients (LCC), colorectal cancer patients (CCC), and diabetes patients (DC). **(B)** Sorting the small RNA category by sequence. Serum and blood cells contain multiple small RNA species (<30 nt) including miRNAs, ribosomal RNA fragments, and mRNA fragments. **(C)** Validation of the Solexa data by semi-qRT-PCR assays with 30 cycles. Behind the name of each serum miRNA is the copy number sequenced by Solexa. Note that serum miRNAs with a copy number lower than 10 are not detected in semi-qRT-PCR assays. **(D)** Number and overlap of miRNAs between female serum (FS) and male serum (MS) samples. **(E)** Pearson correlation scatter plot of miRNA levels between FS and MS ($R=0.8924$). **(F)** Number and overlap of miRNAs between serum and blood cell fraction of healthy donors. **(G)** Comparison of miRNA levels in NS to that in NC ($R=0.9206$).

male subjects (MS), healthy female subjects (FS), and various patients including those with non-small cell lung cancer (NSCLC), colorectal cancer (CCS), and diabetes (DS). As can be seen in this figure, a significant amount of small RNAs in the sera from various sources were 21–23 nt in length, consistent with the ideal size of miRNAs. Similar patterns of small RNAs were also observed with RNAs extracted from blood cells of both healthy subjects (MC and FC) and patients (LCC, CCC, and DC). This result suggests that there is an enriched miRNA fraction in the small RNAs from serum. Figure 2B shows the sorting results of total small RNAs. As shown in the figure, serum and blood cells both contain multiple and heterogeneous small RNA species (<30 nt in length) including miRNAs, rRNA fragments, and mRNA fragments. Compared to blood cells from the same donors, serum contains a relatively lower level of miRNA and a larger proportion of ribosomal RNA fragments. Solexa sequencing results identified approximately 190 known miRNAs in the serum of healthy subjects by referencing a miRNA library (release 10.0).

Expression profile of serum miRNAs in healthy human male and female subjects

We further validated the Solexa results by semi-qRT-PCR using 24 miRNAs with copy number larger than 10 and 9 miRNAs with copy number less than 10. As shown in Figure 2C, results from Solexa and semi-qRT-PCR were quite consistent for those miRNAs with high copy number (>10), but semi-qRT-PCR results from miRNA species with low copy number (<10) were less reliable. Thus, low-copy number miRNAs (<10) were subtracted, which resulted in a profile of 101 miRNAs in healthy Chinese serum (Supplementary information, Table S1). Among these miRNAs, 90 miRNAs were detected in the serum of both male and female subjects, while 10 and 1 miRNAs were only present in the serum of male or female subjects, respectively (Figure 2D). Male-specific serum miRNAs included miR-100, miR-184, and miR-923, while miR-222 represented a female-specific serum miRNA. Judging by the copy number of individual miRNAs in human serum, we found that, for most miRNAs, there is no significant difference in expression level between healthy male and female subjects (Supplementary information, Table S1). As shown in the Pearson correlation scatter plot comparing male serum miRNA to female serum miRNA (Figure 2E), R is very close to 1. Given the nearly identical profiles of healthy male and female subjects, we used normal serum miRNA (NS), the average of miRNA levels in MS and FS, to stand for serum miRNA level in healthy subjects.

Comparison of miRNA expression profile in serum and in blood cellular fraction

We next sequenced all miRNAs in the total RNA samples isolated from blood cells of healthy subjects (NC), and compared this expression profile with that of serum miRNAs. As shown in Figure 2F, most of the miRNAs (91 out of 101) were detected in both serum and blood cells, whereas only a small number of miRNAs were uniquely present in either serum or blood cells. The results suggest that under normal conditions most serum miRNAs are derived from circulating blood cells. The concordance between serum miRNAs and blood cell miRNAs is shown in Figure 2G and detailed in Supplementary information, Table S2.

Expression profiles of serum miRNAs in patients with non-small cell lung carcinoma, colorectal cancer, and type 2 diabetes

To explore the potential of using miRNAs as biomarkers for diseases, we investigated the expression profile of miRNAs in various patients and compared it with that of normal subjects. Employing the Solexa approach, we sequenced whole serum miRNAs of patients with ongoing non-small cell lung carcinoma (NSCLC). As done with normal serum miRNAs, the sera from 11 lung cancer patients were pooled for RNA extraction prior to Solexa analysis. As shown in Figure 3A, the expression profile of miRNA in lung cancer serum (LCS) was significantly different from that of NS. Compared to healthy subjects, 28 miRNAs were missing and 63 new miRNA species were detected in the lung cancer patients. The significant difference in miRNA profiles between NS and LCS samples is summarized in Figure 3B and detailed in Supplementary information, Table S3. As a result, the Pearson correlation coefficient (R) is rather low ($R=0.2429$). Similarly, we sequenced all miRNAs in lung cancer blood cell (LCC) samples and compared them to those in LCS. Surprisingly, the miRNA profile of LCS was also remarkably different from that of LCC. As shown in Figure 3C, 57 miRNAs were shared by LCS and LCC, but 76 miRNAs were detected in LCS only. The differential miRNA expression between serum and blood cells of lung cancer patients is a striking contrast to that of healthy subjects, in which serum and blood cells essentially share the same miRNA profile (Figure 2F). Such difference was illustrated using a Pearson correlation scatter plot (Figure 3D) and is detailed in Supplementary information, Table S4.

To test whether the drastic change in serum miRNAs is specific for NSCLC patients, we performed the Solexa sequencing procedure with serum RNA samples from patients with colorectal cancer and type 2 diabetes. As

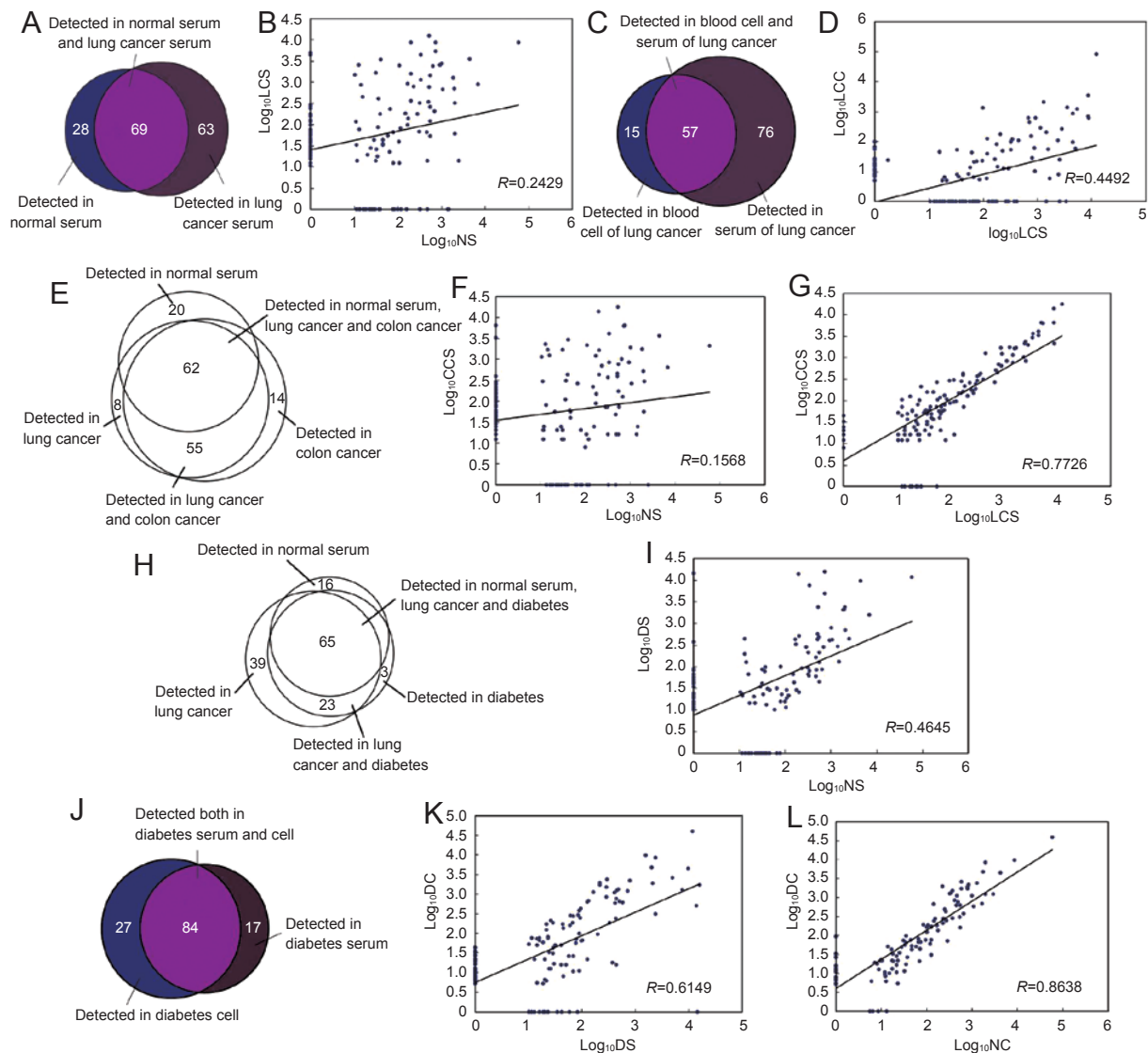


Figure 3 Expression profile of serum miRNA in patients with non-small cell lung carcinoma, colorectal cancer, and type 2 diabetes. (A) Number and overlap of miRNAs between NS and LCS samples. (B) Pearson correlation scatter plot of miRNA levels in NS and LCS ($R=0.2429$). (C) Number and overlap of miRNAs between LCS and LCC samples. (D) Pearson correlation scatter plot of miRNA levels in LCS and LCC ($R=0.4492$). (E) Number and overlap of miRNAs among NS, LCS, and CCS samples. (F) Pearson correlation scatter plot of miRNA levels in NS and CCS ($R=0.1568$). (G) Pearson correlation scatter plot of miRNA levels in LCS and CCS ($R=0.7726$). (H) Number and overlap of miRNAs among NS, LCS, and DS samples. (I) Pearson correlation scatter plot of miRNA levels in NS and DS ($R=0.4645$). (J) Number and overlap of miRNAs between DS and DC samples. (K and I) Pearson correlation scatter plot of miRNA levels in DS and DC, and in NC and DC, respectively.

shown in Figure 3E and 3F, colorectal cancer patients also had a significantly different serum miRNA profile compared to healthy subjects. In all, 69 miRNAs were detected in the colorectal cancer serum (CCS) but not in NS. It is of interest to note that colorectal cancer patients shared a large number of serum miRNAs (e.g., miR-134, miR-146a, miR-221, miR-222, miR-23a, etc.) with lung cancer patients. A Pearson correlation scatter plot further

indicated that the levels of miRNAs in serum from lung cancer patients and colorectal cancer patients were consistent (Figure 3G), suggesting that there are some “common” tumor-related miRNAs in serum.

Compared to healthy subjects, diabetes patients also had a significantly altered expression profile of serum miRNAs, though the change was not as drastic as that in cancer patients (Figure 3H). Surprisingly, diabetes

patients and lung cancer patients share a large number of common serum miRNAs that are not found in healthy subjects. These miRNAs may be related to the body's immune system and alteration of these miRNAs may reflect a general inflammatory response shared by various diseases. A Pearson correlation scatter plot (Figure 3I) confirmed the remarkable difference of miRNAs between diabetes serum (DS) and NS. We have also compared the expression profile of DS miRNAs with that of diabetes blood cell (DC) miRNAs. As shown in Figure 3J, besides 84 common miRNAs shared by DS and DC, there were 17 and 27 miRNAs that were only found in DS and DC, respectively. Pearson correlation scatter plots showed the comparison of miRNAs between DC and DS (Figure 3K), and between DC and NC (Figure 3L), respectively. Since $R_{DC/NC}$ (0.8638) is larger than $R_{DS/NS}$ (0.4645), alteration of serum miRNAs is more sensitive than that of blood cell miRNAs in reflecting the diabetic condition.

Individual validation of Solexa data by qRT-PCR

The Solexa results from pooled serum samples of 11 NSCLC patients were further validated individually by qRT-PCR. Given that miR-25 and miR-223 have been reported to be implicated in tumorigenesis [10], and our Solexa analysis also showing that altered ratios of miR-25 and miR-223 in serum are among the highest ones in LCS samples (the copy numbers of miR-25 in NS and LCS are 739 and 8 662, and miR-223 in NS and LCS are 41 and 3 446, respectively), miR-25 and miR-223 were selected for further individual examination in an independent trial of 152 lung cancer sera (LCS) and 75 normal sera (NS) by qRT-PCR. As shown in Figure 4A and 4B, serum expression levels of miR-25 and miR-223 are significantly increased in LCS than in NS. As for the negative control, serum expression level of let-7a is not different between LCS and NS (Figure 4C). Figure 4D is the summarization of induction fold of miR-25, miR-223 and let-7a in LCS compared to NS. The qRT-PCR results fully support the concept that serum miRNA expression profiles of patients with NSCLC, colorectal cancer, and

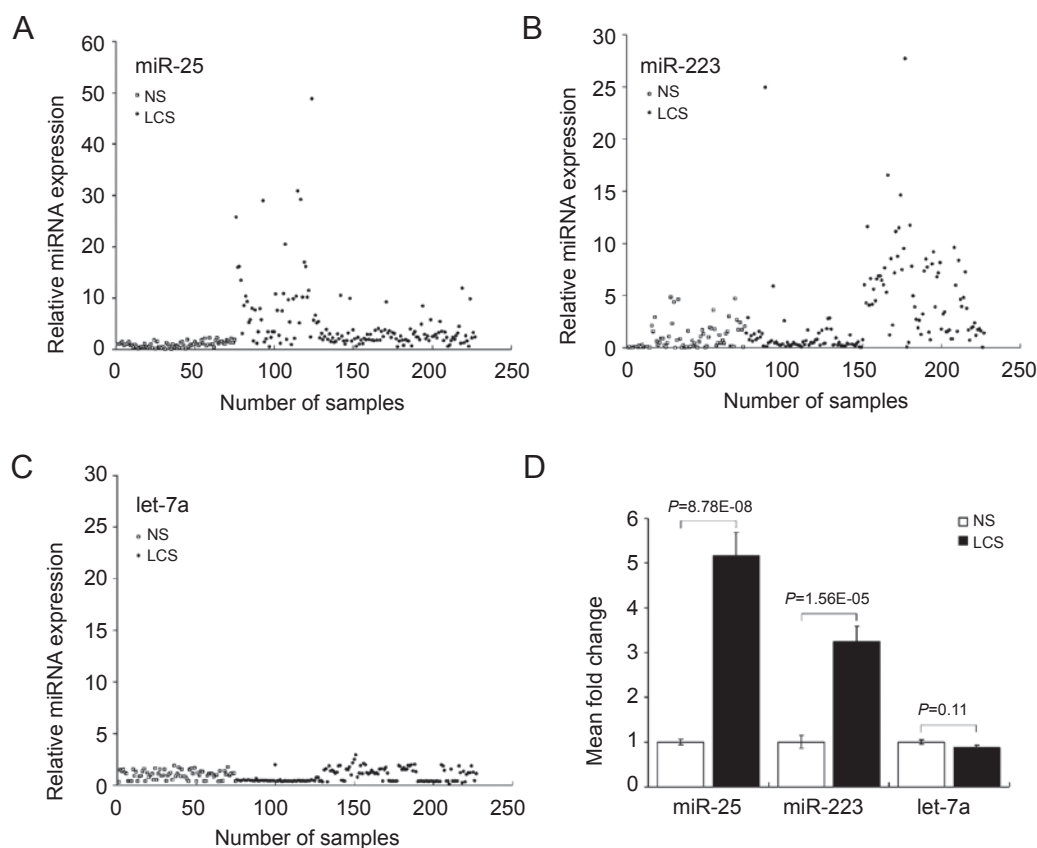


Figure 4 Individual validation of Solexa data by qRT-PCR. (A–C) The expression levels of miR-25 (A), miR-223 (B), and let-7a (C) were measured in 75 NS samples (#1–#75) and 152 LCS samples (#76–#227) by qRT-PCR. The relative miRNA expression level in each sample normalized to the average of NS samples was illustrated. (D) Summarization of the mean fold changes of miR-25, miR-223, and let-7a in LCS compared to NS samples.

diabetes obtained from Solexa analysis are reliable. Also, elevated expression levels of miR-25 and miR-223 in serum are the blood-based biomarkers of NSCLC, which can be easily detected by qRT-PCR.

Discussion

The search for non-invasive tumor markers for diagnosis is currently one of the most rapidly growing areas in cancer research [1-4]. Serum and plasma have been the subject of extensive research for years [2]. However, serum-based test suitable for widespread use in early tumor detection is currently limited [3, 4]. Nowadays, almost all of the routinely used serum markers are proteins and the conventional methodologies used to measure them remain labor-intensive [3, 4]. In the best currently available blood test, carcinoembryonic antigen exhibits low sensitivity and specificity, particularly in the context of early disease. Comprehensive proteomic analysis recently introduced a group of differentially expressed proteins as disease indicators and significantly increased the accuracy of diagnosis [11, 12], but its protein assay procedure is not easy to apply in clinical diagnosis.

The present study is the first one to systematically characterize miRNAs in serum. Our results demonstrate that serum miRNAs are stable and can be detected directly in serum, thereby greatly facilitating clinical use of such tests. Surprisingly, miRNAs, particularly serum miRNAs, are resistant to RNaseA digestion and other harsh conditions, which potentially explains the stability of serum miRNAs. The mechanism of resistance of miRNAs to RNase requires further study. Interestingly, we have found that miRNAs exist not only in human sera/plasma, but also in various animal species. Given that a lot of animals have been used for biomedical research and drug screening, serum miRNA expression profiles of these model animals can potentially serve as useful biomarkers in such studies. Thus, species-specific serum miRNA profiling would be of interest in the future. Furthermore, we have found that miRNAs also exist in other body fluids, including urine, tear, ascetic fluid, and amniotic fluid (data not shown). Obviously, studying miRNA expression profiles in these body fluids would be another important project.

Expression profile of serum miRNA in patients with NSCLC obtained by Solexa analysis shows 63 new miRNAs which are absent in normal subjects. Among these new miRNAs in LCS, many are known to be associated with lung cancer or other tumors. For instance, miR-128b, miR-152, miR-125b, miR-205, miR-27a, miR-146a, miR-222, miR-23a, miR-24, miR-150, *etc.* showed increased levels in tissue samples diagnosed with lung

cancer, while the expression levels of miR-29a, miR-221, miR-223, miR-25, miR-92, miR-99a, *etc.* were increased in colorectal cancer tissues [10]. Moreover, it has been reported that patients with papillary thyroid carcinomas showed increased levels of miR-221, miR-222, and miR-146 in the tumors [13]. Likewise, miR-221, miR-222, and miR-125b were reported to be the top up-regulated miRNAs in TRAIL-resistant non-small cell lung cancer cells [14]. However, among these newly detected serum miRNAs, there are also many, especially those with a high number designation, that have not been previously studied. Since our Solexa results clearly indicate the expression of these miRNAs in serum from cancer patients, their physiological functions and relationship with tumorigenesis should be further examined. The results also strongly suggest that, during diseases such as cancer, serum miRNAs are derived from not only circulating blood cells but also other tissues affected by ongoing diseases, and that these disease-related miRNAs in the serum can serve as potential biomarkers. Furthermore, different miRNA profiles in serum versus blood cells under the disease state again support the conclusion that the serum miRNA profile is not simply a default product of broken blood cells but serves as an indicator of biological function.

Comparing serum miRNA expression profile between patients with non-small cell lung cancer and colorectal cancer, we have found that most of these miRNAs seem to be involved in general tumorigenesis and cell division/growth. Perhaps most interesting is the fact that the Solexa results did identify a unique expression profile of serum miRNAs for each cancer type. For example, 8 and 14 serum miRNAs were uniquely detected in lung cancer and colorectal cancer patients, respectively. For lung cancer patients, the specific miRNAs included miR-205, miR-206, miR-335, *etc.* Some of these miRNAs such as miR-205 have been reported to be significantly up-regulated in bladder cancers [15]. For colorectal cancer patients, those 14 miRNAs are miR-485-5p, miR-361-3p, miR-326, miR-487b, *etc.*

Our present study clearly demonstrates that levels of miRNAs in serum are stable, reproducible, and consistent among individuals of the same animal species. Interestingly, several groups reported that some RNA fragments, including miRNAs, occasionally have been found in serum and plasma, which might be related to certain dysfunctions [16-18]. During the period when this manuscript was previously submitted elsewhere, Lawrie *et al.* [19] reported that miR-21 has the potential as a diagnostic biomarker for diffuse large B-cell lymphoma (DLBCL) and that the sera levels of miR-21 are associated with relapse-free survival in DLBCL patients. Mitchell *et al.*

[20] most recently have found that serum levels of miR-141 can distinguish patients with prostate cancer from healthy controls. These studies, together with our results here, firmly support the notion that miRNAs are stably present in serum or plasma, and could serve as biomarkers for diseases. Moreover, our present study systematically characterized miRNAs in serum. Most importantly, we have identified serum miRNA expression profiles in normal subjects and various diseases. Single or a couple of serum-based biomarkers for one disease, such as AFP for liver cancer, or CRP for inflammation and diabetes *etc.*, have been used for diagnosis over several decades. Lack of sufficient sensitivity, specificity, and accuracy is the limitation of a single blood-based biomarker in clinical use. By contrast, a cluster of biomarkers for one disease would be a better diagnostic tool with much higher sensitivity, specificity, and accuracy. We propose that the specific serum miRNA expression profile (not single or a couple of miRNA(s)) constitutes the fingerprint of a physiological or disease condition, which could have a huge impact on diagnosis and personalized medicine in the future.

In conclusion, we unequivocally show that serum contains large amounts of stable miRNAs derived from various tissues/organs and that the serum miRNA expression profile can be used as a novel serum-based biomarker potentially offering more sensitive and specific tests than those currently available for early diagnosis of cancer and other diseases. This new approach has the potential to revolutionize present clinical management, including determining cancer classification, estimating prognosis, predicting therapeutic efficacy, maintaining surveillance following surgery, as well as forecasting disease recrudescence. Furthermore, given the fact that miRNAs are identified as the first class of RNAs stably present in serum, it would be of great interest for future studies to both understand the biological functions and find other potential applications of serum miRNAs.

Materials and Methods

Serum collection and RNA isolation

Whole blood samples of lung cancer and colorectal cancer were derived from patients at the Tianjin Medical University Cancer Institute and Hospital (Tianjin, China), Jinling Hospital (Nanjing, China), and Shanghai CDC (Shanghai, China) at the time of diagnosis. All of the donors or their guardians provided written consent and Ethics permission was obtained for the use of all samples. Whole blood was separated into serum and cellular fractions within 2 h after blood was derived. Then cellular fractions were immediately frozen in liquid nitrogen and sera were stored at -80°C . For blood cell RNA isolation, total RNA was isolated using Trizol Reagent (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. For serum RNA isolation, equal volume of Trizol was

used, and three steps of phenol/chloroform purification were added since serum is full of proteins. In general, the yield was 5-10 μg RNA/50 ml serum.

Quantitative RT-PCR of mature miRNAs

Assays to quantify the mature miRNAs were conducted as previously described [7, 8] with minor modification. Since U6 and 5S rRNA were degraded in serum samples, and there is no current consensus on the use of house-keeping miRNAs for qRT-PCR analysis (miR-16, one commonly used reference miRNA, is inconsistent in our serum tests), the expression levels of target miRNAs were directly normalized to total RNA. Briefly, serum was purified by using phenol/chloroform extraction to get rid of protein fraction. 20 μl reverse transcriptase reactions contained 10 μl purified serum, 1 \times RT buffer, 0.25 mM each of dNTPs (Takara), 3.33 U/ μl AMV reverse transcriptase (Takara) and 0.25 U/ μl RNase Inhibitor (Takara), and 2 μl antisense looped primer mix. This allowed for the creation of a miRNA cDNA library. The mix was incubated at 16°C for 15 min, 42°C for 60 min, and 85°C for 5 min. Subsequently, real-time quantification was performed using an Applied Biosystems 7300 Sequence Detection system. The 20 μl PCR reaction included 1 μl RT product (1:5 dilution), 0.5 μl Universal reverse primer, 0.5 μl of sense primer, 1 μl SYBR green (Invitrogen), and 1 U/ μl Taq (Takara). The reactions were incubated in a 96-well optical plate at 95°C for 10 min, followed by 40 cycles of 95°C for 15 s and 60°C for 1 min. All reactions were run in triplicate. After reaction, the C_T data were determined using default threshold settings and the mean C_T was determined from the duplicate PCRs. The ratio of cancer serum miRNA to normal serum miRNA was calculated by using the equation $2^{-\Delta\Delta C_T}$, in which $\Delta\Delta C_T = C_{T\text{ cancer}} - C_{T\text{ normal}}$. All primers used are listed in Supplementary information, Table S5.

Solexa sequencing

Briefly, after PAGE purification of small RNA molecules under 30 bases and ligation of a pair of Solexa adaptors to their 5' and 3' ends, the small RNA molecules were amplified using the adaptor primers for 17 cycles and the fragments around 90 bp (small RNA+adaptors) were isolated from agarose gel. The purified DNA was used directly for cluster generation and sequencing analysis using the Illumina's Solexa Sequencer according to the manufacturer's instructions. Then the image files generated by the sequencer were processed to produce digital-quality data. The subsequent procedures performed with Solexa were summarizing data production, evaluating sequencing quality, calculating length distribution of small RNA reads and filtering reads contaminated by rRNA, tRNA, mRNA, snRNA, and snoRNA. Finally, clean reads were compared with a miRBase database (release 10.0) and the total copy number of each sample was normalized to 100 000.

Pearson's correlation coefficient (R)

Correlation is a technique for investigating the relationship between two quantitative, continuous variables. Pearson's correlation coefficient (R), also known as the product-moment coefficient of correlation, is a measure of the strength of the association between the two variables. The first step in studying the relationship between two continuous variables is to draw a scatter plot of the variables to check for linearity. The nearer the scatter of points is to a straight line, the higher the strength of association between the

variables. The Pearson's correlation coefficient (R) may take any value from -1 to $+1$.

Statistical analysis

All photo-images of semi-quantitative RT-PCR were representatives of at least three independent experiments. Quantitative RT-PCR was performed in triplicate, and the entire experiment was repeated several times. Data shown were presented as means \pm SE of three or more independent experiments, and the differences were considered statistically significant at $P < 0.05$ by using the Student's t -test.

Acknowledgments

We thank Drs Fengyong Liu and Sheng Luan at UC Berkeley, USA, for their discussion and help with the writing of the manuscript. This work was supported by grants from the National Natural Science Foundation of China (no. 30225037, 30471991, 30570731), National Basic Research Program of China (973 Program) (no. 2006CB503909, 2004CB518603), the "111" Project, and the Natural Science Foundation of Jiangsu Province (no. BK2004082, BK2006714).

References

- 1 Duffy MJ. Clinical uses of tumor markers: a critical review. *Crit Rev Clin Lab Sci* 2001; **38**:225-262.
- 2 Thomas CM, Sweep CG. Serum tumor markers: past, state of the art, and future. *Int J Biol Markers* 2001; **16**:73-86.
- 3 Duffy MJ. Role of tumor markers in patients with solid cancers: a critical review. *Eur J Intern Med* 2007; **18**:175-184.
- 4 Roulston JE. Limitations of tumour markers in screening. *Br J Surg* 1990; **77**:961-962.
- 5 Esquela-Kerscher A, Slack FJ. Oncomirs – microRNAs with a role in cancer. *Nat Rev Cancer* 2006; **6**:259-269.
- 6 Calin GA, Croce CM. MicroRNA signatures in human cancers. *Nat Rev Cancer* 2006; **6**:857-866.
- 7 Chen C, Ridzon DA, Broomer AJ, *et al.* Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res* 2005; **33**:e179.
- 8 Tang F, Hajkova P, Barton SC, Lao K, Surani MA. MicroRNA expression profiling of single whole embryonic stem cells. *Nucleic Acids Res* 2006; **34**:e9.
- 9 Hafner M, Landgraf P, Ludwig J, *et al.* Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. *Methods* 2008; **44**:3-12.
- 10 Volinia S, Calin GA, Liu CG, *et al.* A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc Natl Acad Sci USA* 2006; **103**:2257-2261.
- 11 Leman ES, Schoen RE, Weissfeld JL, *et al.* Initial analyses of colon cancer-specific antigen (CCSA)-3 and CCSA-4 as colorectal cancer-associated serum markers. *Cancer Res* 2007; **67**:5600-5605.
- 12 Ward DG, Suggett N, Cheng Y, *et al.* Identification of serum biomarkers for colon cancer by proteomic analysis. *Br J Cancer* 2006; **94**:1898-1905.
- 13 He H, Jazdzewski K, Li W, *et al.* The role of microRNA genes in papillary thyroid carcinoma. *Proc Natl Acad Sci USA* 2005; **102**:19075-19080.
- 14 Garofalo M, Quintavalle C, Di Leva G, *et al.* MicroRNA signatures of TRAIL resistance in human non-small cell lung cancer. *Oncogene* 2008; **27**:3845-3855.
- 15 Gottardo F, Liu CG, Ferracin M, *et al.* Micro-RNA profiling in kidney and bladder cancers. *Urol Oncol* 2007; **25**:387-392.
- 16 El-Hefnawy T, Raja S, Kelly L, *et al.* Characterization of amplifiable, circulating RNA in plasma and its potential as a tool for cancer diagnostics. *Clin Chem* 2004; **50**:564-573.
- 17 Tsang JC, Lo YM. Circulating nucleic acids in plasma/serum. *Pathology* 2007; **39**:197-207.
- 18 Chim SS, Shing TK, Hung EC, *et al.* Detection and characterization of placental microRNAs in maternal plasma. *Clin Chem* 2008; **54**:482-490.
- 19 Lawrie CH, Gal S, Dunlop HM, *et al.* Detection of elevated levels of tumor-associated microRNAs in serum of patients with diffuse large B-cell lymphoma. *Br J Haematol* 2008; **141**:672-675.
- 20 Mitchell PS, Parkin RK, Kroh EM, *et al.* Circulating microRNAs as stable blood-based markers for cancer detection. *Proc Natl Acad Sci USA* 2008; **105**:10513-10518.

(Supplementary information is linked to the online version of the paper on the Cell Research website.)