MicroRNA signatures in tissues and plasma predict development and prognosis of computed tomography detected lung cancer

Mattia Boeri[a,1], Carla Verri[a,1], Davide Conte[a,1], Luca Roz[a,1], Piergiorgio Modena[b], Federica Facchinetti[a], Elisa Calabrò[c], Carlo M. Croce[d,2,3], Ugo Pastorino[c,2], and Gabriella Sozzi[a,2,3]

[a]Tumor Genomics Unit, Department of Experimental Oncology and Molecular Medicine, and [c]Unit of Thoracic Surgery, Fondazione IRCCS Istituto Nazionale Tumori, 20133 Milan, Italy; [b]Unit of Experimental Oncology 1, Centro di Riferimento Oncologico, 33081 Aviano (PN), Italy; and [d]Ohio State University Comprehensive Cancer Center, Ohio State University, Columbus, OH 43210

Contributed by Carlo M. Croce, January 12, 2011 (sent for review December 22, 2010)

The efficacy of computed tomography (CT) screening for early lung cancer detection in heavy smokers is currently being tested by a number of randomized trials. Critical issues remain the frequency of unnecessary treatments and impact on mortality, indicating the need for biomarkers of aggressive disease. We explored microRNA (miRNA) expression profiles of lung tumors, normal lung tissues and plasma samples from cases with variable prognosis identified in a completed spiral-CT screening trial with extensive follow-up. miRNA expression patterns significantly distinguished: (*i*) tumors from normal lung tissues, (*ii*) tumor histology and growth rate, (*iii*) clinical outcome, and (*iv*) year of lung cancer CT detection. Interestingly, miRNA profiles in normal lung tissues also displayed remarkable associations with clinical features, suggesting the influence of a permissive microenvironment for tumor development. miRNA expression analyses in plasma samples collected 1–2 y before the onset of disease, at the time of CT detection and in disease-free smokers enrolled in the screening trial, resulted in the generation of miRNA signatures with strong predictive, diagnostic, and prognostic potential (area under the ROC curve ≥ 0.85). These signatures were validated in an independent cohort from a second randomized spiral-CT trial. These results indicate a role for miRNAs in lung tissues and plasma as molecular predictors of lung cancer development and aggressiveness and have theoretical and clinical implication for lung cancer management.

circulating biomarkers | risk prediction | miRNA ratios

Despite recent advances in the management of resected lung cancer and the use of molecular targeted agents in specific clinical settings, the cure rate of non-small-cell lung cancer (NSCLC) remains low due to drug-refractory recurrent and metastatic disease.

Early detection studies using chest X-rays (1) and, more recently, spiral-computed tomography (CT; refs. 2 and 3), have reported a significant increase in the number of lung cancer diagnoses, without apparent major decrease in advanced cancers or reduction of mortality in smokers (4). A recent press release (http://www.cancer.gov) reporting the findings of the largest randomized trial comparing spiral-CT to chest X-rays showed a 6.9% reduction in all-cause mortality (−20.3% lung cancer mortality), but a full report of the results of this trial is not yet available. A likely explanation of the limited impact of CT screening on mortality is that perhaps not all aggressive lung tumors arise from identifiable slow-growing precursors, suggesting a possible paradigm shift in our understanding of the natural history of lung cancer (5, 6). In this respect, the identification of biologic and molecular features of indolent and aggressive disease would be critical to define clinically useful predictors of high-risk lesions.

microRNAs (miRNAs) are small RNA molecules with regulatory function and marked tissue specificity that can modulate multiple targets belonging to several pathways. They are frequently deregulated in cancer (7) and could constitute a new class of blood-based biomarkers useful for cancer detection and prognosis definition because, for their nature, they seem to remain rather intact and stable and are detectable with simple assays like quantitative real-time PCR (qRT-PCR). Initial studies in prostate, colon, and lung cancer patients showed that plasma levels of specific miRNAs had remarkable sensitivity and specificity to distinguish cancer patients from healthy subjects (8–11).

The intent of the present study was to perform an extensive miRNA profiling of primary lung tumors, paired normal lung tissues, and multiple plasma samples collected before and at the time of disease, from two independent spiral CT-screening trials. We aimed to identify biomarkers able to predict tumor development and prognosis, therefore improving lung cancer diagnosis and treatment options.

Results

Lung Cancer Detection and Survival. ***INT-IEO cohort (training set).*** The 5-y screening plan was completed in 2005, and the final results of this screening project were partially published (2, 12).

Lung cancer was diagnosed in 38 subjects, 22 in the first 2 y and 16 from the 3rd to 5th y of screening, including one interval cancer at 4th y. The frequency of stage I was 63% (77% in first 2 y vs. 44% in the last 3 y), and adenocarcinoma was 71% (95% in first 2 y vs. 63% in the last 3 y; Table S1). Median follow-up time for the 38 lung cancer cases was 75 mo, with 60% 5-y overall survival (95% C.I.: 43–74%). Five-y overall survival was 92% for stage I and 7% for stages II–IV ($P < 0.001$; Fig. 1*A*). When the year of detection was considered, 5-y overall survival was 77% for cancers diagnosed in the first 2 y compared with 36% for those detected from 3rd to 5th y of screening ($P = 0.005$; Fig. 1*B*), indicating that incident cancers represent a more aggressive disease. Year of detection and tumor stage were significantly associated (χ^2 test, $P = 0.034$). In the subset of CT year 1–2/stage I, 5-y survival was 94% (95% C.I.: 65.0–99.1). In the whole group of stage I, after exclusion of one death from second primary lung cancer and one from end-stage chronic obstructive pulmonary disorder (COPD), 5-y survival was 100%.

Multicentric Italian Lung Detection (MILD) cohort (validation set). At the end of 4th year of screening in the MILD trial, lung cancer was diagnosed in 53 subjects, 24 in the first 2 y, and 23 in the 3rd and 4th year. Six interval cancers were diagnosed: one in the 1st y, two in the 2nd y, and three in the 3rd y. Early stage disease (Ia–Ib) was diagnosed in 28 (53%) patients, and adenocarcinoma was di-

Author contributions: M.B., C.V., L.R., C.M.C., U.P., and G.S. designed research; M.B., C.V., D.C., and F.F. performed research; M.B., C.V., D.C., L.R., P.M., E.C., U.P., and G.S. analyzed data; and M.B., C.V., L.R., C.M.C., and G.S. wrote the paper.

The authors declare no conflict of interest.

[1]M.B., C.V., D.C., and L.R. contributed equally to this work.

[2]C.M.C., U.P., and G.S. contributed equally to this work.

[3]To whom correspondence may be addressed. E-mail: carlo.croce@osumc.edu or gabriella.sozzi@istitutotumori.mi.it.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1100048108/-/DCSupplemental.

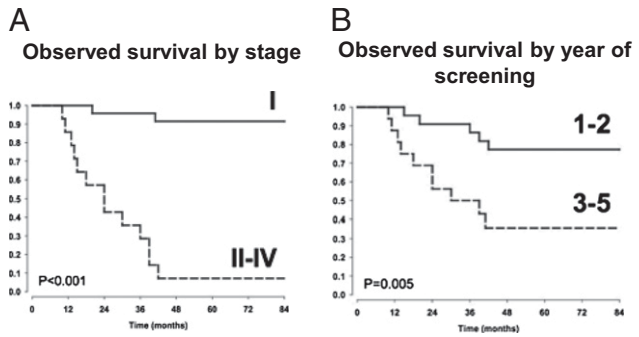

Fig. 1. Kaplan–Meier estimates of observed 5-y survival in CT-screening INT-IEO trial. (*A*) Data arranged according to the extent of disease: 92% for stage I (95% CI: 70.0–97.8) and 7% for stage II–IV (95% CI: 0.5–27.5, $P < 0.001$). (*B*) Data arranged according to the year of CT-detection: 77% for lung cancers detected in the first 2 y of the study (95% CI: 53.7–89.8) and 36% for lung cancers diagnosed from third to fifth years (95% CI: 13.7–58.7, $P = 0.005$)

agnosed in 30 (57%) of patients. Because this trial is ongoing, no interim analysis was performed so far. However, even if the median follow-up time of 23 mo is relatively short, we could divide the 53 patients in two groups of reasonable size: 14 patients with poor prognosis (dead or alive with incurable disease) and 39 patients with good prognosis (alive without disease).

miRNA Expression Profiling in Tumor and Normal Lung. miRNA profiles of 28 tumors and 24 paired normal lung tissues were analyzed using a miRNA microarray platform. Validation of the differentially expressed miRNAs was done using qRT-PCR.

By class comparison and class prediction analyses (using both paired and unpaired algorithms), expression of 56 miRNAs was significantly different at the nominal 0.001 level of the univariate test. The top 10 deregulated miRNAs that discriminate CT-detected lung cancer from normal lung tissue were: mir-7, mir-21, mir-200b, mir-210, mir-219-1, miR-324 (up-regulated), mir-126, mir-451, mir-30a, and mir-486 (down-regulated; Table 1). This list included alterations previously identified in symptomatic lung cancer patients (e.g., mir-21 and the mir-200 family, known to be involved in pathways such as survival, apoptosis, epithelial-mesenchymal transition) and some unidentified changes (e.g., down-regulation of miR-486 and miR-451).

To validate the results obtained with microarray hybridization, the levels of the two most regulated miRNAs (mir-21 and mir-486) were evaluated in tumor and normal samples by qRT-PCR, which confirmed the previous observation.

miRNA Expression in Tissues Is Associated with Clinical-Pathological Features. Possible association of miRNA expression profiles with clinical–pathological characteristics of the patients was then investigated (Table 2). Two miRNAs (mir-205 and mir-21) significantly discriminated adenocarcinoma from squamous cell carcinoma histotypes ($P \leq 0.001$). Mir-518e and mir-144 were down-regulated in tumors with a faster growth rate, and higher levels of mir-429, member of the mir-200 family, correlated with a worse disease-free survival (DFS).

Table 1. Top 10 miRNAs deregulated between tumor and normal lung tissue (class comparison analysis)

	Tumor vs. normal tissues	
miRNAs deregulated ($P < 0.001$)	Direction	Fold change
mir-7-2-prec	Up	1.3
mir-126	Down	0.4
mir-200b	Up	1.3
mir-210	Up	3
mir-219-1	Up	1.6
mir-21	Up	2.9
mir-324-5p	Up	1.3
mir-451	Down	0.5
mir-486-5p	Down	0.5
mir-30a	Down	0.6

The miRNA expression profile of tumors detected in the first 2 y of the screening was significantly different from the profile of tumors appearing after the 2nd y, with differential expression of eight miRNAs (mir-128, mir-129, mir-369-3p, mir-193, mir-339-3p, mir-185, mir-346, and mir-340). These results indicate that these groups of tumors display different miRNA profiles associated with distinct aggressive features, where the incident tumors grow faster.

Interestingly, miRNA expression analysis on normal lung tissues also discriminated subjects identified in the first 2 y from those of later years of screening (miR-126*, mir-126, let-7c, mir-222, mir-30e, mir-1-2, mir-29b-1, mir-30d-prec, mir-15a, mir-16; Fig. 2). Significant associations were found between miRNAs expression in normal lung and reduction of forced expiratory volume (FEV; mir-379 and mir-29-1*), faster tumor growth (mir-30d*), DFS of the patients (mir-34b; Table 2). The results obtained by microarray hybridization were independently validated by qRT-PCR.

Interestingly, although there was no significant difference in smoking habits (packs-per-year, time from smoking cessation), patients detected in years 3–5 showed a higher proportion of severe COPD (GOLD criteria ≥ 2, 33% vs. 5%; χ^2 test, $P = 0.02$).

These findings indicate that specific miRNA signatures in normal lung microenvironment are associated with tumor aggressiveness and clinical history of the patients.

Pathways Enrichment Analysis. For the miRNA signature discriminating tumor from normal samples, pathway enrichment analysis was performed using DIANA-mirPath software on the gene targets predicted by microT-4.0, Pic-Tar, and TargetScan-5. This analysis showed that many of the predicted miRNA targets are involved in critical pathway affected in cancer such as survival, apoptosis, epithelial–mesenchymal transition, and proliferation (Table S2).

miRNA Expression Profiling In Plasma Samples: Study Design. Validated circulating biomarkers in plasma/serum could potentially represent the gold standard for a noninvasive routine clinical application. We reasoned that ideal miRNA biomarkers should be identified before the onset of the tumors and be able to predict aggressive versus indolent disease development.

To determine whether specific miRNA signatures are already detectable in plasma samples collected before the detection of the disease, we performed high-throughput miRNA expression profiles of plasma samples using TaqMan microfluidic cards (Applied Biosystems). We first analyzed plasma samples collected >1 y before disease development and at the time of disease detection (positive CT/surgery) in the training set (CT-screening trial INT-IEO). We generated miRNA signatures that were then validated in plasma samples (also predisease and at disease detection) of a validation set (CT-screening MILD cohort). The clinical–pathological characteristics of training and validation sets are shown in Table S3. As control groups, we tested 15 pools of plasma samples (5–7 individuals per pool, 81 individuals in total) collected from disease-free subjects (negative spiral-CT) from both trials, with age, sex, and smoking habits distribution similar to those of cases.

Using microfluidic cards, 113 miRNAs were found to be always expressed in all plasma samples, and a subset of 100 miRNAs was found to be consistently expressed in the 15 control pools, with a good reproducibility among biological duplicates (Fig. S1). These 100 miRNAs were then used to identify circulating biomarkers of risk, diagnosis, and prognosis in plasma samples collected before or in presence of CT-detected disease.

Table 2. Associations between miRNA expression in tumor and normal tissues and clinical–pathological characteristics of patients

	Tumor tissue			Normal tissue		
Clinical–pathological characteristics	miRNA	Direction	*P* value	miRNA	Direction	*P* value
Histotype (ADC vs. SCC or others)	mir-205	Down	<0.001			
	mir-21-pre	Up	<0.001			
Growth rate diameter (≤50% vs. >50%)	mir-518e	Up	<0.001	mir-30d*	Up	<0.001
	mir-144-pre	Up	<0.001			
Disease-free survival (alive vs. dead or relapse)	mir-429	Down	0.003	mir-34b	Up	0.001

miRNA Ratios as Bioinformatics Tools for miRNA Analysis. Because the normalization of miRNA data in plasma samples is still a controversial issue, the ratios between the expression values of all miRNAs consistently expressed in plasma were computed. Each value of a single miRNA was compared with the values of all of the other 99 miRNAs, and 4,950 ratios were obtained and subsequently used to analyze differences between classes of samples resulting in the definition of ratios with clinical relevance (*Materials and Methods*). When using microfluidic cards, there is general agreement on the normalization of single miRNA expression using the mean values of expression of all miRNAs of each card (13). To validate the robustness of the miRNA ratios method, we compared the results obtained independently by the two methods in the microfluidic cards. The results showed that the miRNAs mostly deregulated in multiple ratios were the same as those detected using the normalization on the mean expression value, thus confirming the robustness of the ratios method.

The use of miRNA ratios seems to be an easily applicable method with potential for general clinical use that avoids the need for large scale, high-throughput analyses and was therefore used to develop clinically useful signatures based on circulating biomarkers.

Identification of Diagnostic and Prognostic Circulating miRNA Profiles in Plasma Samples Collected Before and at the Time of Disease Detection. Class comparison analysis was initially performed in the training set to identify a group of miRNA ratios showing statistically significant differences between prediagnostic, diagnostic, and disease-free plasma ($P < 0.05$). These ratios were then technically validated, in a subset of samples, by TaqMan MicroRNA assays.

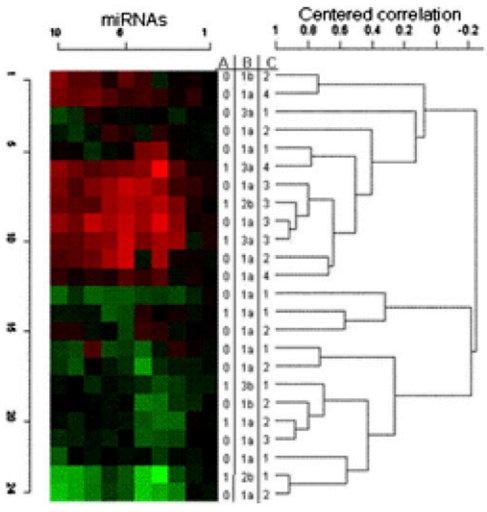

Fig. 2. Clustering analysis on 24 normal lung tissue samples using miRNAs differentially expressed between patients with tumors detected in the first 2 y and those of later years of screening. Clinical status of the patient (0 = alive, 1 = dead), tumor stage, and year of tumor detection are reported in columns A, B, and C, respectively.

To assess the consistency of miRNA ratios within the control pools, we compared the value of each ratio in two control pools with the mean value resulting from the analysis of the individual samples composing the pools. We found that the values were consistent.

However, because some ratios showed a high individual variability in the control subjects, possibly leading to a high number of false positives, we considered for further analyses only those ratios with a minimal intrapool variability.

The signatures obtained were then used to calculate specificity and sensitivity in an independent validation set.

Because the range of miRNA expression levels in the two datasets was consistently different, possibly due to a storage effect (14), the patients in each dataset were compared with the respective control groups.

For the generation of the signatures predicting clinical outcome (both before and in presence of CT-detected disease), because of the small number of events, we grouped the two datasets. Cases with bad outcome were compared with the respective control pools, and the signatures obtained were then tested for their power to discriminate patients with bad (dead and alive with disease) or good (disease free) prognosis in the whole cohort.

miRNA signature identifies individuals at risk to develop lung cancer. To investigate whether there are molecular markers predicting development of lung cancer, samples collected from patients 1 and/or 2 y before detection of the disease by CT were analyzed and compared with the control pools of heavy-smoking individuals (Fig. 3 *A* and *B*).

A signature of 16 ratios composed by 15 miRNAs could discriminate correctly 18 of 20 samples from subjects developing lung cancer in the training set (90% sensitivity) and resulted positive in only 1 of the 5 control pools (80% specificity). In the validation set, this signature identified 12 of 15 samples collected before lung cancer detection by spiral-CT, with sensitivity of 80%

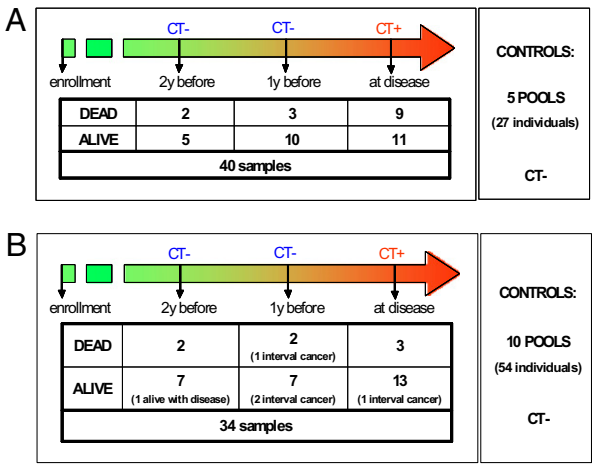

Fig. 3. Diagram of samples collection and analysis in the training set (INTIEO trial; *A*) and in the validation set (MILD trial; *B*).

and specificity of 90% (AUC-ROC = 0.85, $P < 0.0001$; Fig. 4*A*). The predictive value of this signature was evaluated to be useful up to 28 mo before the disease, and mir-660, mir-140-5p, mir 451, mir-28-3p, mir-30c, and mir-92a are the most frequently deregulated miRNAs.

miRNA signature with diagnostic value. Plasma samples collected at surgery or at time of disease detection by spiral CT were compared with pools of disease-free individuals to identify a miRNA profile associated with lung cancer diagnosis. In the training set, a panel of 16 ratios involving 13 different miRNAs classified 16 of 19 patients, with a sensitivity of 84% and a specificity of 80%. In the validation set plasma samples, 12 of 16 patients were correctly discriminated, with a sensitivity of 75% and a specificity of 100% (AUC-ROC = 0.88, $P < 0.0001$; Fig. 4*B*).

The lower sensitivity observed may be related to the presence of a higher number of small, early-stage nodules with indolent behavior in this series and the inclusion of two patients misclassified by both the signature of diagnosis and risk.

The diagnostic signature was then used for class prediction of predisease plasma samples in the same series. In the training set, 11 of 20 (55%) cases were classified as individuals with disease and, very interestingly, 10 of these 11 cases were characterized by poor prognosis (dead or alive with disease) or belonged to the group of patients identified from 3rd to 5th y of screening. In the validation set, similar results were obtained, with presence of the disease signature already in 10 of 15 (66.6%) predisease plasma samples. Moreover, looking at the three predisease samples of interval cancer cases (patients who developed lung cancer few months after a negative CT result), only 1 patient was classified by the risk signature. Instead, 2 cases (including the 1 identified by risk signature) already displayed the diagnostic signature 8–9 mo before disease detection. The interval cancer case not recognized by any signatures had a stage 1a tumor with good outcome, suggesting the presence of a low-risk nodule.

Only 4 ratios were shared by the signatures of risk and of diagnosis, and the miRNAs involved were partially different. mir-17, mir-660, mir-92a, mir-106a, and mir-19b were the most frequently deregulated at the time of lung cancer diagnosis.

Overall, these findings strengthen the observation that circulating miRNA in plasma is detectable well before clinical disease detection by spiral CT, indicating the possibility to select high-risk groups on the basis of miRNA profiling.

miRNA signature of risk to develop aggressive lung cancer. We analyzed the miRNA expression profiles in predisease plasma samples of individuals with poor clinical outcome to define a signature of miRNAs identifying individuals at high risk to develop an aggressive disease.

A signature of 10 ratios, composed of 9 different miRNAs, identified 5 of 5 patients with poor prognosis (dead or with progressive disease) in this first set (100% sensitivity and 100% specificity). In the validation set, 4 of 5 patients with poor prognosis were correctly classified, including a patient with poor prognosis who developed an interval cancer. The sensitivity of this signature in the validation set was 80% with 100% specificity.

mir-221, mir-660, mir-486-5p, mir-28-3p, mir-197, mir-106a, mir-451, mir-140-5p, and mir-16 are the miRNAs deregulated in the signature of aggressive disease.

The signature was then used for class prediction of predisease plasma samples of patients with good prognosis in training and validation sets. The signature identified 5 of 15 (33.3%) patients

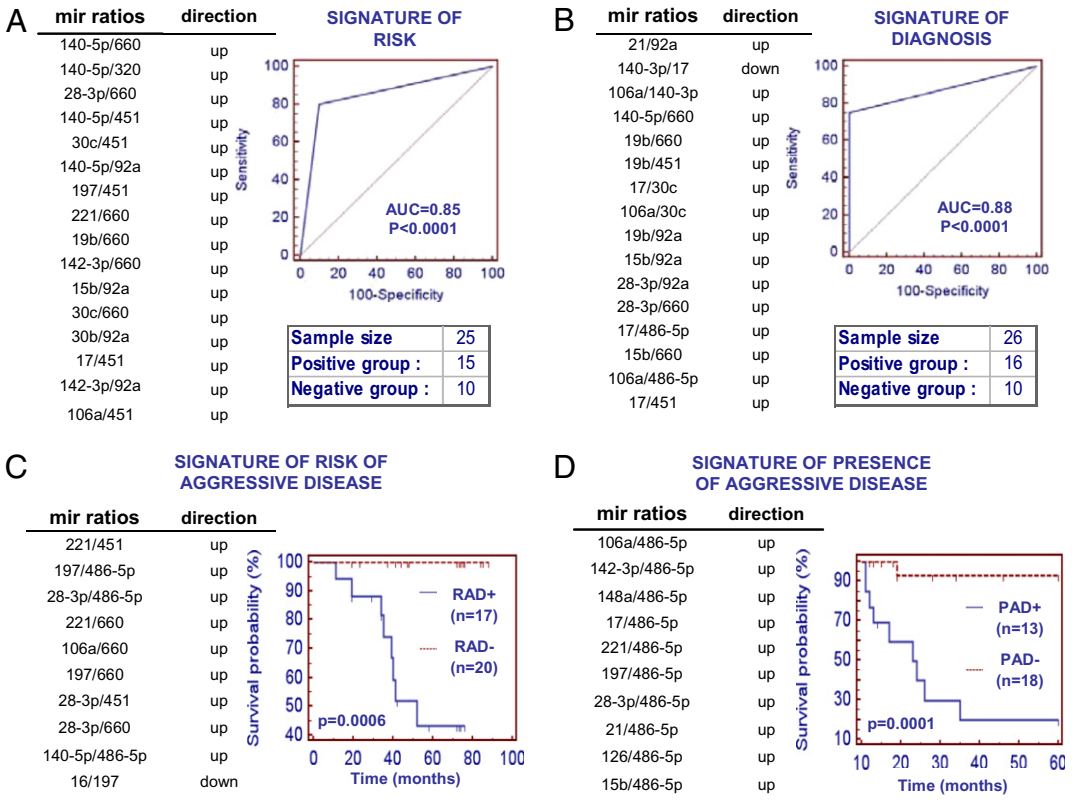

Fig. 4. miRNA expression analyses in plasma samples collected before the onset and at the time of disease. The signatures of miRNA ratios and their direction in the analyses are listed in the tables. (*A*) miRNA signature of risk to develop lung cancer and (*B*) miRNA signature of lung cancer diagnosis. The ROC curves of samples belonging to the validation set are shown. (*C*) Kaplan–Meier survival curves of patients with miRNA signatures of risk of aggressive disease (RAD) in plasma samples collected 1–2 y before CT-detection of lung cancer. (*D*) Kaplan–Meier survival curves of patients with miRNA signatures of presence of aggressive disease (PAD) in plasma samples collected at the time of CT-detected lung cancer. The RAD- or PAD-positive patients show a significantly worse survival rate than RAD- or PAD-negative patients ($P = 0.0006$ and $P = 0.0001$, respectively).

in the training set and 5 of 11 (45%) patients in the validation set (Fig. 4*C*). Interestingly, in both the datasets, most of these classified samples belonged to patients whose tumor was detected after the 3rd y of screening. This finding supports our previous observation on tissue samples where a distinct miRNA profile was identified in tumor and normal tissues of the same patients. Noticeably, among the patients with tumor diagnosed in the 2nd y of screening (all stage Ia and Ib tumors), only one case with stage 1b tumor had the risk signature of aggressive disease.

These results suggest that miRNA profiles in predisease plasma samples are able to predict the development of tumors with worse prognosis and might even be helpful in pinpointing those early stage tumors at high risk of aggressive evolution.

miRNA expression in plasma samples at time of disease detection and prognosis. Then we looked at the association between miRNA expression and prognosis in plasma samples collected at the time of lung cancer diagnosis by generating a signature composed by 10 ratios, all containing mir-486-5p. This signature identified 7 of 8 patients with bad prognosis in the training set (88% sensitivity and 100% specificity). The signature of aggressive disease was observed also in 2 of 10 samples with good prognosis, one of these having a stage Ib tumor. In the validation set, only 3 plasma samples collected in presence of disease of patients with poor prognosis were available, and 2 of these had the profile of aggressive disease. The third case was misclassified by all of the analyses performed in all plasma samples collected during screening evaluations (Fig. 4*D*).

Again, this signature was used for class prediction of predisease plasma samples of patients in the training and validation sets. Half of the predisease samples of patients with bad prognosis were positive for both the signatures of aggressive disease, whereas the predisease samples of patients with good prognosis that showed the signature of aggressive disease belonged mainly (5 of 6) to patients with tumors detected after the 3rd y of screening. It is noteworthy that, although individuals in the training set have an extended follow-up and 5-y overall survival data are available, the shorter median follow-up observation time (14 mo) for patients in validation set might affect the strength of the prognostic signatures.

mir-486-5p, compared with mir-21, mir-126, mir-15b, mir-148a, mir-142-3p, mir-17, mir-197, mir-221, mir-28-3p, and mir-106a, appears to be always down-regulated in plasma of patients with bad outcome.

Discussion

The investigation of biological and molecular features of indolent and aggressive lung cancer is critical to identify specific risk markers for lung cancer development, to achieve the earliest possible prediction and intervention and, potentially, to define novel therapeutic targets.

In this study, we have focused on the role of miRNAs as biomarkers of lung disease by taking advantage of the availability of both tissue samples (tumor and normal lung) and multiple plasma samples, collected before and at the time of disease detection, from patients enrolled in two different spiral-CT screening trials with extended follow-up. These patients developed tumors displaying variable aggressive behavior during the course of the trials.

Although previous studies reported miRNA expression profiles predicting recurrence and prognosis only in lung tumor samples collected at the time of surgery for symptomatic lung cancer (15–17), our study provides unique results on miRNA signatures able to identify the presence of aggressive lung cancer not only in tumor, but also in normal lung tissues and in plasma samples of patients. Moreover, miRNAs deregulated in plasma samples collected before clinical appearance of disease were powerful molecular predictors of high-risk disease development.

In tumor samples, we confirmed up-regulation of known miRNAs such as mir-21, a miRNA with proproliferative and anti-apoptotic function that is reported to target PTEN (18), and described down-regulation of two miRNAs (mir-486 and mir-451) that are involved in maintenance of self-renewal capacity of bronchio-alveolar stem cells (19). Association analyses revealed that expression of mir-205 and mir-21 are markers linked to squamous cell carcinoma (SCC) and adenocarcinoma (ADC) histology, respectively, confirming previous studies on the validity of studying miRNA expression in support of histopathological diagnosis for a precise classification of tumor histology (20–22). Interestingly, miRNAs that were deregulated in the more aggressive tumors identified in later years of screening are involved in adhesion and invasion pathways: miR-339 was reported to negatively regulate intercellular cell adhesion molecule (ICAM)-1 (23), and mir-128a has been involved in TGFβ pathway promotion of tumor cell invasion and metastasis (24). This miRNA specifically targets FOXO1A, a transcription factor involved in AKT signaling and apoptosis inhibition (25).

The finding of miRNA expression profiles associated with aggressive disease and poor survival in normal lung tissues of the patients strengthens the existing evidence on the critical influence of the normal lung microenvironment on tumor development and, in the present study, on tumor aggressiveness. It is possible to speculate that these markers might represent molecular signs of a "soil" that, after extensive damage caused by smoking, becomes permissive, or even promoting, for cancer development. Several miRNAs deregulated in normal lung tissue of the patients undergoing surgery are involved in major pathways linked to cancer. In particular, miR-126 is known to promote angiogenesis by repressing the inhibitors of VEGF signaling spred1 and pik3r2 (26), and let-7 is involved in proinflammatory programs (27). In addition, AKT signaling is the major pathway influenced by miR-222 (28), miR-30 regulates connective tissue growth factor (29), and mir-29b modulates anti-apoptotic and prometastatic matrix molecules by repressing Mcl-1 (30). It is also interesting to note the down-regulation of mir-34b in patients with worse DFS, because mir-34b is a well known target of p53, which cooperates to control cell proliferation and adhesion-independent growth (31). The observation of a possible prognostic role of several miRNAs in normal lung opens up the possibility of innovative therapeutic strategies targeting the host rather than the tumor itself.

Because circulating miRNAs in plasma could be more tissue-specific than tumor-specific, we decided to perform a high-throughput miRNA expression in plasma profiling using microfluidic cards. We then developed multiplex real-time PCR assays to validate, as single PCR assays, those miRNA signatures significantly associated with clinical characteristics of the patients. We have optimized simple and highly reproducible miRNA assays and formulated a suitable algorithm for qRT-PCR data validation in plasma using miRNA reciprocal ratios. Our findings suggest that the assessment of a number of miRNAs in plasma by qRT-PCR assays is a potentially useful and clinically applicable procedure to improve lung cancer management.

miRNAs deregulated in tissue specimens were rarely detected in plasma samples, further strengthening the high tissue-specificity of miRNAs and suggesting a predictive role of plasma miRNAs independent from tissue specimens. We observed that a partially different set of miRNAs were deregulated in plasma before and at the time of disease. This finding might be explained by the consideration that genes and pathways necessary in the earlier phases of disease development are different from those required for the maintenance and the progression of the tumor.

Overall, the 21 miRNAs composing the signatures of risk, diagnosis, and prognosis in plasma belong to major pathways: cellular aging (mir-19b, mir-17, mir-106; ref. 32), bronchioalveolar and hematopoietic stem cells renewal (mir-486, mir-106a, 142-3p; ref. 19), tumor recurrence in stage I NSCLC (mir-27b; mir-106a; mir-19b; mir-15b mir-16, mi-21; refs. 16 and 33), and lung cancer aggressiveness (mir-221, mir-222; ref. 34). In particular mir-17, mir-92a, mir-19b, and mir-106a are oncomirs belonging to the same family responsible for increased proliferation, repression of apoptosis. and induction of angiogenesis (35, 36). mir-197 regulates expression of the tumor suppressor gene FUS1, whose expression is lost in a large proportion of lung tumors. mir-28-3p is located in a chromosomal region that is

frequently amplified in lung cancer (3q28; ref. 37). mir-221 blocks PTEN expression leading to activation of the AKT pathway, and is suggested to play an important role in cell growth and invasiveness by targeting the PTEN/AKT pathway. Alterations of these pathways represent well established and meaningful risk factors in lung cancer. Finally, in a recent publication regarding circulating miRNAs, mir-21, mir-126, and mir-486-5p were also identified as potential blood-based biomarkers with diagnostic value in NSCLC patients (38).

The identification of miRNA signatures in plasma samples collected 1–2 y before disease that predict cancer development and prognosis is potentially useful in the selection of high-risk individuals who need to undergo spiral-CT surveillance. It is noteworthy that specific miRNA signatures in predisease plasma samples are able to predict and discriminate the development of the more aggressive, early metastatic tumors that are frequently undetectable by yearly spiral-CT. This information could be certainly helpful to prompt these individuals in pharmacological smoking cessation programs and possibly to propose more specific imaging for detection of occult metastatic disease (e.g., PET, whole-body MRI), as well as nontoxic treatments such as enrollment in prophylactic vaccination programs. Furthermore, the signature of a potentially aggressive disease could also help in the clinical management of the frequent early-stage nodules detected during CT-screening trials improving diagnostic algorithms.

Considering the noninvasive characteristics of plasma sampling and the reproducible and easy detection of miRNA markers, we envision that plasma-based miRNA biomarkers could be used in clinical practice and may help to avoid overdiagnosis and overtreatment of low-risk disease and late detection of high-risk and early metastatic disease.

Materials and Methods

Study Populations, CT Screening Protocols, and Samples Collection. Recruitment of high-risk population and diagnostic imaging workup have already been described (12, 39) and are reported in detail in *SI Materials and Methods*.

miRNA Profiling in Tissues and Plasma Samples. Isolation of total RNA and miRNA expression profiling are described in detail in *SI Materials and Methods*.

Statistical and Bioinformatics Analyses. Microarray and qRT-PCR data handling and analyses are described in detail in *SI Materials and Methods*.

Analysis of miRNA ratios in plasma samples. The *Ct* value of each miRNA obtained with SDS 2.2.2 (Applied Biosystems) was transformed in the corresponding expression value (2^{-Ct}). We then calculated the ratios between miRNAs, considering only one ratio for each pair of miRNAs. For the ratios showing statistically significant differences in the class comparison analyses ($P < 0.05$), a cutoff value was established with the formula (mean of the ratio in one class $\pm$ SD) to use the ratios as binomial variables in the generation of different signatures. The control pools from the technical validation of the training set were used to establish the optimal number of features of each signature to discriminate patients in the validation set.

ACKNOWLEDGMENTS. This work was supported by grants from Associazione Italiana Ricerca Cancro (AIRC) (to G.S. and U.P.); the Italian Ministry of Health (Ricerca Finalizzata, P.I.O.) (to G.S.); Fondazione Ermenegildo Zegna (Biella, Italy), Fondazione Cariplo and Regione Lombardia (Milan, Italy), and Early Detection Research Network Program of the National Cancer Institute Grant UO1Ca152758 (to C.M.C.).

1. Fontana RS, et al. (1986) Lung cancer screening: the Mayo program. *J Occup Med* 28: 746–750.
2. Bach PB, et al. (2007) Computed tomography screening and lung cancer outcomes. *JAMA* 297:953–961.
3. Infante M, et al.; DANTE Study Group (2009) A randomized study of lung cancer screening with spiral computed tomography: three-year results from the DANTE trial. *Am J Respir Crit Care Med* 180:445–453.
4. Pastorino U (2010) Lung cancer screening. *Br J Cancer* 102:1681–1686.
5. Bach PB (2008) Is our natural-history model of lung cancer wrong? *Lancet Oncol* 9: 693–697.
6. Chirieac LR, Flieder DB (2010) High-resolution computed tomography screening for lung cancer: unexpected findings and new controversies regarding adenocarcinogenesis. *Arch Pathol Lab Med* 134:41–48.
7. Esquela-Kerscher A, Slack FJ (2006) Oncomirs - microRNAs with a role in cancer. *Nat Rev Cancer* 6:259–269.
8. Mitchell PS, et al. (2008) Circulating microRNAs as stable blood-based markers for cancer detection. *Proc Natl Acad Sci USA* 105:10513–10518.
9. Ng EK, et al. (2009) Differential expression of microRNAs in plasma of colorectal cancer patients: A potential marker for colorectal cancer screening. *Gut* 58:1375–1381.
10. Chen X, et al. (2008) Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. *Cell Res* 18:997–1006.
11. Hu Z, et al. (2010) Serum microRNA signatures identified in a genome-wide serum microRNA expression profiling predict survival of non-small-cell lung cancer. *J Clin Oncol* 28:1721–1726.
12. Pastorino U, et al. (2003) Early lung-cancer detection with spiral CT and positron emission tomography in heavy smokers: 2-year results. *Lancet* 362:593–597.
13. Mestdagh P, et al. (2009) A novel and universal method for microRNA RT-qPCR data normalization. *Genome Biol* 10:R64.
14. Sozzi G, et al. (2005) Effects of prolonged storage of whole plasma or isolated plasma DNA on the results of circulating DNA quantification assays. *J Natl Cancer Inst* 97: 1848–1850.
15. Yanaihara N, et al. (2006) Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. *Cancer Cell* 9:189–198.
16. Patnaik SK, Kannisto E, Knudsen S, Yendamuri S (2010) Evaluation of microRNA expression profiles that may predict recurrence of localized stage I non-small cell lung cancer after surgical resection. *Cancer Res* 70:36–45.
17. Landi MT, et al. (2010) MicroRNA expression differentiates histology and predicts survival of lung cancer. *Clin Cancer Res* 16:430–441.
18. Zhang JG, et al. (2010) MicroRNA-21 (miR-21) represses tumor suppressor PTEN and promotes growth and invasion in non-small cell lung cancer (NSCLC). *Clin Chim Acta* 411:846–852.
19. Qian S, et al. (2008) MicroRNA expression profile of bronchioalveolar stem cells from mouse lung. *Biochem Biophys Res Commun* 377:668–673.
20. Lebanony D, et al. (2009) Diagnostic assay based on hsa-miR-205 expression distinguishes squamous from nonsquamous non-small-cell lung carcinoma. *J Clin Oncol* 27:2030–2037.
21. Markou A, et al. (2008) Prognostic value of mature microRNA-21 and microRNA-205 overexpression in non-small cell lung cancer by quantitative real-time RT-PCR. *Clin Chem* 54:1696–1704.
22. Bishop JA, et al. (2010) Accurate classification of non-small cell lung carcinoma using a novel microRNA-based approach. *Clin Cancer Res* 16:610–619.
23. Ueda R, et al. (2009) Dicer-regulated microRNAs 222 and 339 promote resistance of cancer cells to cytotoxic T-lymphocytes by down-regulation of ICAM-1. *Proc Natl Acad Sci USA* 106:10746–10751.
24. Masri S, et al. (2010) The role of microRNA-128a in regulating TGFbeta signaling in letrozole-resistant breast cancer cells. *Breast Cancer Res Treat* 124:89–99.
25. Myatt SS, et al. (2010) Definition of microRNAs that repress expression of the tumor suppressor gene FOXO1 in endometrial cancer. *Cancer Res* 70:367–377.
26. Fish JE, et al. (2008) miR-126 regulates angiogenic signaling and vascular integrity. *Dev Cell* 15:272–284.
27. Polikepahad S, et al. (2010) Pro-inflammatory role for let-7 microRNAs in experimental asthma. *J Biol Chem* 285:30139–30149.
28. Wong QW, et al. (2010) MiR-222 overexpression confers cell migratory advantages in hepatocellular carcinoma through enhancing AKT signaling. *Clin Cancer Res* 16: 867–875.
29. Duisters RF, et al. (2009) miR-133 and miR-30 regulate connective tissue growth factor: implications for a role of microRNAs in myocardial matrix remodeling. *Circ Res* 104:170–178.
30. Mott JL, Kobayashi S, Bronk SF, Gores GJ (2007) mir-29 regulates Mcl-1 protein expression and apoptosis. *Oncogene* 26:6133–6140.
31. Corney DC, Flesken-Nikitin A, Godwin AK, Wang W, Nikitin AY (2007) MicroRNA-34b and MicroRNA-34c are targets of p53 and cooperate in control of cell proliferation and adhesion-independent growth. *Cancer Res* 67:8433–8438.
32. Hackl M, et al. (2010) miR-17, miR-19b, miR-20a, and miR-106a are down-regulated in human aging. *Aging Cell* 9:291–296.
33. Duncavage E, Goodgame B, Sezhiyan A, Govindan R, Pfeifer J (2010) Use of microRNA expression levels to predict outcomes in resected stage I non-small cell lung cancer. *J Thorac Oncol* 5:1755–1763.
34. Garofalo M, et al. (2009) miR-221&222 regulate TRAIL resistance and enhance tumorigenicity through PTEN and TIMP3 downregulation. *Cancer Cell* 16:498–509.
35. Matsubara H, et al. (2007) Apoptosis induction by antisense oligonucleotides against miR-17-5p and miR-20a in lung cancers overexpressing miR-17-92. *Oncogene* 26: 6099–6105.
36. Osada H, Takahashi T (2011) let-7 and miR-17-92: small-sized major players in lung cancer development. *Cancer Sci* 102:9–17.
37. Du L, et al. (2009) miR-93, miR-98, and miR-197 regulate expression of tumor suppressor gene FUS1. *Mol Cancer Res* 7:1234–1243.
38. Shen J, et al. (November 29, 2010) Plasma microRNAs as potential biomarkers for non-small-cell lung cancer. *Lab Invest*, 10.1038/labinvest.2010.194.
39. Pastorino U, et al. (2009) Fluorodeoxyglucose uptake measured by positron emission tomography and standardized uptake value predicts long-term survival of CT screening detected lung cancer in heavy smokers. *J Thorac Oncol* 4:1352–1356.