INDEX

INTRODUCTION

Lung cancers represent the first cause of oncological deaths in the world and their global incidence and mortality rates are rising [1]. Despite major discoveries in tumor biology and molecular genomics [2, 3], targeted therapies can only be delivered to a small number of individuals (i.e. *EGFR* and *ALK* inhibitors) [4,5]. Likewise, therapeutic options remain inadequate and reliable or reproducible markers are needed to effectively stratify lung cancer patients and improve their clinical outcomes [6,7].

Histologically, lung cancers are classically subdivided in Small-Cell Lung Cancer (SCLC) and Non Small-Cell Lung Cancer (NSCLC) and this stratification has relevant clinical and therapeutic implications. Among NSCLC the most prevalent tumor entity is represented by Adenocarcinoma (ADC, 38% of all lung cancer) [8]. Although all ADC share common histological features [8], they include a heterogeneous collection of different lesions. Yet, the pathological stratification of NSCLC, even if parceled, guides clinical and molecular decisions; for example, *EGFR* mutations are investigated solely in case of lung ADC [9].

As highlighted by the 2011 guidelines of IASLC/ATS/ERS [10] and by the new WHO classification of lung tumors [11], ADC include multiple tumor types. These can be further defined based on the expression of selected markers and unique molecular lesions. Since the integration of histological and molecular classifiers has been successfully proven to dissect specific clinical-biological entities among many human neoplasms (i.e. lymphoid tumors), it is reasonable that an analogous method can also improve the stratification of lung cancers. This should eventually lead to an objectively applicable classification based on a new taxonomy.

We hypothesized that ADC retain features related to their cell origin. Thus, we selected a set of markers preferentially expressed by alveolar [Thyroid Transcription Factor 1 (TTF-1), Surfactant Protein A (SP-A) and Napsin A] and bronchiolar [Mucin 5AC (MUC5AC), Caudal Type Homeobox 2 (CDX-2) and Cytokeratin 5 (CK5)] cells to further stratify lung ADC. Neoplasms were defined based on the differentiation profiles and then correlated with *EGFR, KRAS, PIK3CA, BRAF, NRAS* somatic mutations [3].

Our data demonstrate that either integrated morphological stratification or the genetic characterization are not strong predictors of patients' outcomes. The integration of these parameters with the profiles defined using cell of origin markers, improves the clinical prediction of lung ADC patients. This opens new modalities for the design and implementation of patient's tailored therapies.

MATERIALS AND METHODS

PATIENTS COHORT AND CLINICOPATHOLOGICAL FEATURES

Patients from 2003 to 2013 who underwent surgical resection with curative intent for lung adenocarcinoma at Department of Thoracic Surgery - Azienda Ospedaliero Universitaria Città della Salute e della Scienza di Torino, Italy, as well as patients treated with preoperative protocols (i.e. chemotherapy, radiotherapy) were retrospectively reviewed. Routine clinical formalin-fixed paraffin embedded (FFPE) blocks and archived tissue samples within the Institutional Biobank (San Giovanni Battista Hospital, University of Torino, Italy) were examined. [12].

Age, gender, smoking habits, vascular invasion, tumor grading, pTNM staging data and the incidence of previous malignancies were collected.

PATHOLOGICAL EVALUATION

Cases were reviewed by expert pathologists (M.C., L.D., A.N. and F.M.) and subdivided into 6 subtypes, according to the 2015 WHO classification of lung tumors [9]. Samples were stratified based on the most represented morphological patterns as follow: 1.lepidic, 2.acinar, 3.papillary, 4.micropapillary, 5.solid and 6.mucinous.

TISSUE MICRO-ARRAY AND IMMUNOPHENOTYPIC ANALYSES

Three different cores (0.5 mm) with enriched tumor content (>50%) were selected from each tumor sample and multi-tissue microarrays were constructed (Beecher Instruments, Inc., Silver Spring, MD, USA). Serial (4-μm-thick) sections were used for immuno-histochemical analyses, and staining was performed using a semi-automated instrument [13, 14]. Tissue sections were incubated with the following primary antibodies: TTF1 (clone: 8G7G3/1, source: Dako, Carpinteria, CA, USA, dilution: 1:50), Napsin A (clone: TMU-Ad02, source: ARP, Waltham, MA, USA, dilution: 1:500), SP-A (clone: 32E12, source: Novocastra, Burlingame, CA, USA, dilution 1:200), MUC5AC (clone: CLH2, source: Novocastra, dilution: 1:100), CDX2 (clone: CDX2-88, source: Biogenex, Fremont, CA, USA, dilution: 1:40) and CK5 (clone: XM26, source: Novocastra, dilution: 1:100). All samples were processed using a sensitive ‘Bond polymer Refine’ detection system in an automated Bond immunohistochemistry instrument (Vision-Biosystem, Leica, Milan, Italy). Negative control sections were generated in absence of a primary antibody.

The percentage of positive neoplastic cells was recorded for each case, and a compiled score was assigned following the subsequent criteria: 0 to +1 value to the alveolar markers and a 0 to -1 value to the bronchiolar markers (see Table 1). Based on the expression patterns of the 6 markers panel, four different immunophenotypes were defined: alveolar-type (expressing mainly alveolar markers: TTF1, Napsin 1, SP-A), bronchiolar-type (expressing mainly bronchiolar markers: MUC5AC, CDX2, CK5), mixed-type (expressing both of them) and null-type (all markers were negative) (Fig. 1).

GENOMIC DNA EXTRACTION AND HOT SPOT MUTATIONS

Genomic DNAs were extracted from FFPE sections (5 μm thick) with high tumor content (>50%) with the Maxwell® 16 FFPE Tissue LEV DNA Purification Kit (Promega) and total DNA quantified by NanoDrop 2000 (Thermo scientific). Somatic mutations were detected by the Myriapod® Cancer Status system (Diatech Pharmagenomics) and the Sequenom MassARRAY iPlex platform (Sequenom, USA) (see Table Suppl. 1).

STATISTICAL ANALYSIS

The characteristics of patients were recorded and then compared with histological pattern, mutation status of selected genes and immunophenotypical classes. Categorical data are presented as number (percentage, %), and continuous data as median [interquartile range, (IQR)] or as mean with standard deviation (SD).

Between-group differences in patients' characteristics were evaluated by the Mann–Whitney U test for continuous variables and the chi-square test or Fisher's exact test for categorical variables.

The overall survival (OS), estimated with the Kaplan–Meier method [15], was defined as the time from the date of intervention to the date of death by any cause; patients still alive were censored on the date of last follow-up.

A Cox proportional hazard model was employed to estimate the crude and adjusted hazard ratios (HRs) and 95% and to evaluate possible predictors of survival.

The proportional hazard assumption was also verified by graphical checks and formal tests based on Schoenfeld residuals. As the proportional hazard assumption was not met for the histological patterns ($p=0.014$), given the relative homogeneity of effect of the two histological subtypes on OS and in order to have a more consistent reference group, the acinar and the papillary tumors were grouped.

The HRs for each histological pattern (introduced into the model as dummy variables, considering the acinar/papillary pattern as the reference group), for the mutation status (reference group: wild type mutation) and for immunophenotype classes (reference group: alveolar differentiation) were analyzed in separated models, and adjusted for several baseline characteristics (age, sex, smoking status, time since surgery, presence of previous malignancies, presence of vascular invasion and stage of the tumor) known to be predict patients survival.

After careful evaluation and exclusion of strong collinearity between the variables considered, the final model takes into account all these characteristics on the OS.

Statistical significance ($p<0.05$) of differences in OS among periods of surgery, gender, age at diagnosis and tumor types was tested using the log-rank test for homogeneity [16, 17]. As a sensitivity analysis, the same models were replicated excluding patients lacking the immuno-phenotype characterization (N=68).

Statistical analyses were performed using the Stata 13.1 software (StataCorp LP, College Station, TX, USA).

RESULTS

According to the criteria of selection, a total of 531 lung adenocarcinomas were included. Two hundred and seventy-two (51.3%) patients were alive at the time of data collection and 258 (48.7%) died during the extended follow-up period. The patients' mean age was 66.5 years (SD ±8.6); most of the patients were male (67%) and 338 ever smokers (41.1% and 22.6% former and current smokers, respectively). The median patients follow-up time was 3.55 years (IQR 1.96-6.71).

Table 2 shows the distribution of main clinicopathological variables among the predominant histological pattern [8]. Acinar or papillary adenocarcinoma subtypes were the most frequent (60.7%), while only 29 (5.5%) and 17 (3.2%) cases with mucinous and micropapillary adenocarcinoma were reported. Sex, smoking status, tumor grading, pTNM stage, and visceral pleural invasion were significantly associated with predominant histological pattern. A statistical association between these groups was reported for hot spot mutations status ($p<0.001$) and for distinct immunophenotype classes ($p=0.021$).

The distribution of the same clinicopathological characteristics of patients by the different immunophenotype classes is described in Table 3. The most common subset was represented by tumors with alveolar differentiation pattern (279 patients - 52.5%) followed by neoplasms with bronchiolar differentiation or alveolar and bronchiolar markers expression (mixed type) (both 84 patients - 15.8%). Meanwhile tumors of "null-type" category represented a small minority (16 patients - 3%).

Alveolar-differentiated adenocarcinomas were more commonly seen in young ($p=0.06$) female ($p=0.08$) patients, frequently harboring *EGFR* mutations ($p=0.003$). Moreover, alveolar differentiation was associated with acinar/papillary histological pattern ($p<0.001$). Conversely, bronchiolar-differentiated adenocarcinomas were found in older individuals, preferentially male and they associated with a higher degree of vascular invasion ($p=0.01$) and frequent mutation of *KRAS* genes ($p=0.07$). Bronchiolar differentiation resulted linked with mucinous and solid histological pattern ($p<0.001$). No associations between predominant immuno-phenotype class and smoking status, previous malignancies, tumor grading and pTNM stage were found. Analysis of prognostic factors for OS were reported in Table 4 and Fig 2. At univariate analysis, age at diagnosis ($p=0.003$), males ($p=0.028$), pTNM stage of tumor ($p<0.001$) and presence of vascular invasion ($p<0.001$) were significant prognostic factors for OS. Patients with mucinous (HR=1.96, CI95%: 1.20-3.21) and solid histological pattern (HR=1.52, CI95%: 1.13-2.03) had a significant increased risk of deaths when compared to patients with the acinar/papillary neoplasms (Table 4).

Survival rates among acinar/papillary patterns were higher (5-years survival rate was 54.2) as compared to those of solid and mucinous subtypes (5-years survival rates were 17.2% and 28.0% respectively) (Fig 2A).

Different somatic mutations did not proved to be a significant predictor for OS both at univariate and at multivariate analysis. Nevertheless patients presenting PI3K mutations appear to have the worst survival, even if because of the small numbers of cases did not reach a statistical significance (HR 1.61, $p= 0.148$) (Table 4). Survival rates by different somatic mutations showed similar trends

among *EGFR* mutated, *KRAS* mutated and Wild Type patients; 5-year survival rates range from 25.0% in patients with PI3K mutations to 38.7% in patients *KRAS* mutated ADC (Fig 2B).

By univariate analysis mixed (HR 1.45 CI95%04-2.02, p= 0.028), bronchiolar (HR 1.60 CI95% 1.14-2.25, p= 0.006), and null-type differentiated (HR 2.38 CI95% 1.34-4.22, p= 0.003) subsets displayed significantly reduced survival rates (Table 4). Five-years survival rates of patients stratified using cell-of-origin markers, were 37.6%, 27.7% and 39.3% for alveolar, bronchiolar and mixed type differentiation, respectively (Fig. 2C).

Table 4 also illustrated the final multivariate model, which encompasses the histological pattern, hot spot mutations status and the immunophenotype. Using this approach, patients' age (HR=1.03 CI95%: 1.01-1.05), pTNM stage (Stage IIA-IIB HR=1.96 CI95%: 1.38-2.76; IIIA-IIIB HR=3.23 CI95%: 2.28-4.58; IV HR=4.38 CI95%: 2.89-6.63) and the vascular invasion (HR=1.51 CI95%: 1.13-2.03) were confirmed to be good clinical predictors in lung ADC. Among different methods analyzed, final multivariate Cox model showed that the immunophenotype of neoplastic cells was the only independent predictor of patients' outcomes.

CONCLUSIONS

We identify a 6 “cell of origin” markers panel achievable through a routine IHC platform enables the sub-stratification of these tumors in different prognostic groups. Notably, bronchiolar- and null-type appeared to have a negative effect on patients’ prognosis. Even when compared with morphology appearance or mutational status, cell-based origin immunophenotype remains, at a multivariate analysis, the solely parameter able to predict survival; thus highlighting the prognostic value of this improved lung ADC taxonomy.

At a multivariate analysis the clinical value of histologic pattern has not been confirmed. As a single variable, solid predominant and specifically mucinous invasive tumors showed a worse prognosis. This finding is in line with the pre-existing literautre [18-25].

Next, we evaluated the presence of genetic defects on “hot spot” genes frequently mutated in lung ADC [3]. We reported mutational ranges (*KRAS* 35%, *EGFR* 16%, *PI3K* 3%) similar those previously reported by others [26]. Interestingly, *EGFR* mutations were associated more frequently to the lepidic histology and the alveolar phenotype, conversely, *KRAS* to mucinous histology and bronchiolar phenotype as well as *PI3K* mutations. None of the mutations predict patients’ survival.

Different attempts have been conducted to ameliorate lung adenocarcinoma classification and different strategies proposed: a grading system (similar to one adopted in prostate tumors) evaluating metastasizing properties based on the major morphologic subtype [27], five-cell type classification coupled with gene mutations (*p53, KRAS, EGFR*) [28,29], TTF1 expression and gene expression strategies. These latter were used to evaluate the entire cell transcriptome [30] or define specific “high risk” classifiers (as like airway Basal Cell signature) [31, 32].

With the purpose to correlate lung ADC to their histogenesis and underpinning pathogenetic events, we evaluated our dataset with a restricted antibodies panel to identify a hypothetic “bronchiolar” or “alveolar” derivation, and link these data to molecular events and potentiate patients’ survival stratification.

The specific contribution of the various cell types of the airway epithelium to lung carcinogenesis and tumor heterogeneity has not yet well understood. Airway Basal Cells (BC) are hypothesized to be the cell of origin of lung squamous cell carcinoma, considering that airway BC are likely the source of potential pre-neoplastic lesions (i.e. squamous cell metaplasia) [33].

Conversely, the cellular origin of lung ADC has not been clearly identified. Indeed, if centrally located pulmonary adenocarcinoma are thought to arise from the surface or glandular epithelium of bronchi, peripheral adenocarcinoma could arise from Clara cells and type II pneumocytes, since they can exhibit such differentiation pattern, also known as “terminal respiratory unit”. Furthermore, subpopulations of Clara cells have been related to lung adenocarcinoma development in murine models. Lastly, the contribution of other types of cell, such as BC-like progenitors, to human lung adenocarcinoma has been proposed [34, 35, 36].

Therefore, based on differential embriogenetic derivation, we designed a panel of

alveolar and bronchiolar immuno-histochemical markers, routinely available in pathology suites.

Thyroid transcription factor 1 (TTF-1) is a transcriptional factor of pneumocytes type II that plays a crucial role in epithelial morphogenesis stimulating protein synthesis of surfactant. It represents a marker of alveolar and terminal respiratory unit and is expressed in lung adenocarcinoma. Napsin A is an aspartic proteinase that is present in type II pneumocytes, an useful marker in the differential diagnosis of lung ADC. Apoproteins A of surfactant (SP-A) are lung-specific proteins, which play an important role in diminishing superficial tension of the alveoli, mainly localized in type II pneumocytes. TTF-1, Napsin A and SP-A are alveolar markers. MUC-5AC is a secretory mucin with strongly and diffusely expression in mucinous type bronchiole-alveolar carcinomas. CDX-2 is a transcription factor involved in the regulation of intestinal epithelial cells and results positive in the majority of intestinal adenocarcinomas. However, different studies demonstrated that CDX-2 is expressed also in intestinal type adenocarcinomas of the lung, so called “enteric type” lung adenocarcinoma [37]. CK-5 is a high-weight cytokeratin, and its expression has been observed in a minority of lung adenocarcinomas, with a scattered positivity, as observed in our series. MUC-5AC, CDX-2 and CK5 represent the bronchiolar markers.

Based on these 6-markers, we identify 2 major categories: “alveolar type” adenocarcinoma expressing type II pneumocytes markers (TTF-1, SP-A1 and Napsin A) and “bronchiolar type” adenocarcinoma expressing goblet cells markers (MUC5AC, CK5 and CDX2). The first group derives from putative bronchioalveolar stem cell precursors at the bronchioalveolar duct junctions [35] and, consistently, these tumors harbored more often *EGFR* mutations previously reported in studies which have evaluated the Terminal Respiratory type of lung ADC [38, 39]. The latter type comprises adenocarcinoma expressing mucinous cells features and originates from airway Basal Cells of bronchi or bronchioles; these neoplasms express more frequently *KRAS* mutations [40, 41, 42, 43]. Of note, in these tumors the lack of TTF1/Nkx2-1 expression, a master regulator of pulmonary differentiation, is associated with the appearance of a latent gastric differentiation program (evaluated by CDX2 expression) normally restrained by TTF1, as demonstrated experimentally by murine models [44].

A consistent number of tumors expressed markers of both derivation (“mixed” group) and some ADC lack a diffuse or strong expression, defining a “null” type, which is characterized by the worst clinical behavior.

This classification is supported also by embryological considerations; the lung epithelium derives from the foregut endoderm following a developmental-related diversification into peripheral and proximal cell lineages. Consequently the histological appearance of lung carcinoma is extraordinarily diverse and this heterogeneity is consistent with a variety of genetic alterations affecting different epithelial cell precursors homing the lung tissue. For example, a peculiar subtype of pulmonary ADC, called enteric adenocarcinoma, retains the expression of intestinal marker (CDX2, CK20, MUC2 and Villin) and demonstrates a bronchiolar immunoprofile (with loss or reduced expression of the alveolar markers TTF-1, Napsin A and SP-A) [37]. Thus, the possible involvement of endodermal stem cells retaining multiple differentiation potentials could be hypothesized in the development of these tumors.

The usage of a limited panel of 6 markers is predicted to become very useful during routine practice, to stratify lung ADC and guide molecular analysis. Alveolar and bronchiolar types seems to harbor different genetic mutations, and this may be used to reliably predict which molecular tests should be executed, in absence of strong indication based on pure morphology [45].

The lack of strong prognostic markers for lung ADC urges to evaluate different strategies. Our classification goes further a single parametric stratification; the immune-phenotype based on cell-of-origin markers has broader implication being subjective to molecular disrupting events, differentiation commitments and/or morphologic appearance; it allows to identify different lung ADC subtypes efficiently predicting patients' survival. Indeed, at multivariate analysis, this cell-of-origin immunophenotype remains the only independent predictor of survival, also when compared to histology and genetics.

In conclusion, in the absence of clinical prognostic markers, the 6 "cell of origin" markers' classifier is 1) more predictable than morphologic parameters and the genetic fingerprint 2) represents an independent predictor of survival in a multivariate analysis and 3) opens new perspectives in the study and classification of lung adenocarcinoma.

REFERENCES

1. Siegel RL, Miller KD, Jemal A: Cancer Statistics, 2015. CA Cancer J Clin 65:5-29, 2015.
2. Cancer Genome Atlas Research Network et al: Comprehensive genomic characterization of squamous cell lung cancers. Nature 489:519-25, 2012.
3. Cancer Genome Atlas Research Network et al: Comprehensive molecular profiling of lung adenocarcinoma. Nature 511:543-50, 2014.
4. Russo A, Franchina T, Ricciardi GR, et al: A decade of EGFR inhibition in EGFR-mutated non small cell lung cancer (NSCLC): Old successes and future perspectives. Oncotarget 6:26814-25, 2015.
5. Shaw AT, Kim DW, Nakagawa K, et al: Crizotinib versus Chemotherapy in Advanced ALK-Positive Lung Cancer. N Engl J Med 368:2385-94, 2013.
6. Krawczyk P, Kucharczyk T, Kowalski DM, et al: Polymorphisms in TS, MTHFR and ERCC1 genes as predictive markers in first-line platinum and pemetrexed therapy in NSCLC patients. J Cancer Res Clin Oncol 140:2047-57, 2014.
7. Naidoo J, Drilon A: KRAS-Mutant Lung Cancers in the Era of Targeted Therapy. Adv Exp Med Biol 893:155-78, 2016.
8. Travis WD: Pathology of Lung Cancer. Clin Chest Med 32:669-92, 2011.
9. Maemondo M, Inoue A, Kobayashi K, et al: Gefitinib or chemotherapy for non-small-cell lung cancer with mutated EGFR. N Engl J Med 362:2380-8, 2010.
10. Travis WD, Brambilla E, Noguchi M, et al: International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma. J Thorac Oncol 6:244-85, 2011.
11. Travis WD, Brambilla E, Nicholson AG, et al: The 2015 World Health Organization Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances Since the 2004 Classification. J Thorac Oncol 10:1243-60, 2015.
12. Guerrera F, Tabbò F, Bessone L, et al: The Influence of Tissue Ischemia Time on RNA Integrity and Patient-Derived Xenografts (PDX) Engraftment Rate in a Non-Small Cell Lung Cancer (NSCLC) Biobank. PLoS One 11:e0145100, 2016.
13. Luchini C, Parcesepe P, Nottegar A, et al: CD71 in Gestational Pathology: A Versatile Immunohistochemical Marker With New Possible Applications. Appl Immunohistochem Mol Morphol 24:215-20, 2016.

14. Luchini C, Parcesepe P, Mafficini A, et al: Specific expression patterns of epithelial to mesenchymal transition factors in gestational molar disease. Placenta 36:1318-24, 2015.
15. Kaplan EL, Meier P: Non-parametric estimation from incomplete observations. J Am Stat Assoc 53:457-81, 1958.
16. Fisher LD, Van Belle G: Biostatistics, a methodology for the health sciences (ed 10). New York, John Wiley & Sons, 1993, pp 801-3.
17. Peto R, Pike MC, Armitage P, et al: Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. analysis and examples. Br J Cancer 35:1-39, 1997.
18. Russell PA, Wainer Z, Wright GM, et al: Does lung adenocarcinoma subtype predict patient survival? A clinicopathologic study based on the new International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society international multidisciplinary lung adenocarcinoma classification. J Thorac Oncol 6:1496-504, 2011.
19. Yoshizawa A, Motoi N, Riely GJ, et al: Impact of proposed IASLC/ATS/ERS classification of lung adenocarcinoma: prognostic subgroups and implications for further revision of staging based on analysis of 514 stage I cases. Mod Pathol 24:653-64, 2011.
20. Hung JJ, Yeh YC, Jeng WJ, et al: Predictive value of the international association for the study of lung cancer/American Thoracic Society/European Respiratory Society classification of lung adenocarcinoma in tumor recurrence and patient survival. J Clin Oncol 32:2357-64, 2014.
21. Murakami S, Ito H, Tsubokawa N, et al: Prognostic value of the new IASLC/ATS/ERS classification of clinical stage IA lung adenocarcinoma. Lung Cancer 90:199-204, 2015.
22. Xu CH, Wang W, Wei Y, et al: Prognostic value of the new International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society classification in stage IB lung adenocarcinoma. Eur J Surg Oncol. 41:1430-6, 2015.
23. Warth A, Muley T, Meister M, et al: The novel histologic International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society classification system of lung adenocarcinoma is a stage-independent predictor of survival. J Clin Oncol 30:1438-46, 2012.

24. Zhao ZR, Xi SY, Li W, et al: Prognostic impact of pattern-based grading system by the new IASLC/ATS/ERS classification in Asian patients with stage I lung adenocarcinoma. Lung Cancer 90:604-9, 2015.
25. Hung JJ, Yeh YC, Jeng WJ, et al: Prognostic Factors of Survival after Recurrence in Patients with Resected Lung Adenocarcinoma. J Thorac Oncol 10:1328-36, 2015.
26. Meador CB, Micheel CM, Levy MA, et al: Beyond histology: translating tumor genotypes into clinically effective targeted therapies. Clin Cancer Res 20:2264-75, 2014.
27. Sica G, Yoshizawa A, Sima CS, et al: A grading system of lung adenocarcinomas based on histologic pattern is predictive of disease recurrence in stage I tumors. Am J Surg Pathol 34:1155-62, 2010.
28. Hashimoto T, Tokuchi Y, Hayashi M, et al: Different subtypes of human lung adenocarcinoma caused by different etiological factors. Evidence from p53 mutational spectra. Am J Pathol 157:2133-41, 2000.
29. Okada A, Shimmyo T, Hashimoto T, et al: Predictive advantage of a cell type classification for pulmonary denocarcinoma coupled with data for p53, K-ras and EGFR alterations. Cancer Sci 101:1745-53, 2010.
30. Takeuchi T, Tomida S, Yatabe Y, et al: Expression profile-defined classification of lung adenocarcinoma shows close relationship with underlying major genetic changes and clinicopathologic behaviors. J Clin Oncol 24:1679-88, 2006.
31. Beer DG, Kardia SL, Huang CC, et al: Gene-expression profiles predict survival of patients with lung adenocarcinoma. Nat Med. 8:816-24, 2002.
32. Bhattacharjee A, Richards WG, Staunton J, et al: Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. Proc Natl Acad Sci U S A 98:13790-5, 2001.
33. Rock JR, Onaitis MW, Rawlins EL, et al: Basal cells as stem cells of the mouse trachea and human airway epithelium. Proc Natl Acad Sci U S A 106:12771-5, 2009.
34. Fukui T, Shaykhiev R, Agosto-Perez F, et al: Lung adenocarcinoma subtypes based on expression of human airway basal cell genes. Eur Respir J 42:1332-44, 2013.
35. Kim CF, Jackson EL, Woolfenden AE, et al: Identification of bronchioalveolar stem cells in normal lung and lung cancer. Cell 121:823-35, 2005.

36. Ooi AT, Mah V, Nickerson DW, et al: Presence of a putative tumor-initiating progenitor cell population predicts poor prognosis in smokers with non-small cell lung cancer. Cancer Res 70:6639-48, 2010.
37. Nottegar A, Tabbò F, Luchini C, et al: Pulmonary adenocarcinoma with enteric differentiation: immunohistochemistry and molecular morphology. Appl Immunohistochem Mol Morphol 2016 Oct 7. [Epub ahead of print]
38. Yatabe Y, Kosaka T, Takahashi T, et al: EGFR mutation is specific for terminal respiratory unit type adenocarcinoma. Am J Surg Pathol 29:633-9, 2005.
39. Takeuchi T, Tomida S, Yatabe Y, et al: Expression profile-defined classification of lung adenocarcinoma shows close relationship with underlying major genetic changes and clinicopathologic behaviors. J Clin Oncol 24:1679-88, 2006.
40. Marchetti A, Buttitta F, Pellegrini S, et al: Bronchioloalveolar lung carcinomas: K-ras mutations are constant events in the mucinous. J Pathol 179:254-9, 1996.
41. Yoshizawa A, Sumiyoshi S, Sonobe M, et al: Validation of the IASLC/ATS/ERS lung adenocarcinoma classification for prognosis and association with EGFR and KRAS gene mutations: analysis of 440 Japanese patients. J Thorac Oncol 8:52-61, 2013.
42. Kadota K, Yeh YC, D'Angelo SP, et al: Associations between mutations and histologic patterns of mucin in lung adenocarcinoma: invasive mucinous pattern and extracellular mucin are associated with KRAS mutation. Am J Surg Pathol 38:1118–1127, 2014.
43. Cho HC, Lai CY, Shao LE, et al: Identification of tumorigenic cells in Kras(G12D)-induced lung adenocarcinoma. Cancer Res 71:7250-8, 2011
44. Snyder EL, Watanabe H, Magendantz M, et al: Nkx2-1 Represses a Latent Gastric Differentiation Program in Lung Adenocarcinoma. Mol Cell 50:185-99, 2013.
45. Inamura K, Ninomiya H, Ishikawa Y, et al. Is the Epidermal Growth Factor Receptor Status in Lung Cancers Reflected in Clinicopathologic Features? Arch Pathol Lab Med 134:66-72, 2010.

TABLES

Table 1. The scoring-system calculated for the classification of the pulmonary adenocarcinomas as alveolar, bronchiolar, mixed (both alveolar and bronchiolar markers expression, final sum of 0) and null type (no expression of neither alveolar nor bronchiolar markers, final sum of 0).

Antibody of alveolar differentiation	Score 0	Score +1
TTF-1	Expression in ≤25% of neoplastic cells	Expression in >26% of neoplastic cells
Napsin A	Expression in ≤25% of neoplastic cells	Expression in >26% of neoplastic cells
SP-A1	Negative	Expression in ≥1% of neoplastic cells
Antibody of bronchiolar differentiation	Score 0	Score -1
MUC5AC	Negative	Expression in ≥1% of neoplastic cells
CK5	Negative	Expression in ≥1% of neoplastic cells
CDX2	Negative	Expression in ≥1% of neoplastic cells

Table 2. Characteristics of 531 patients with lung adenocarcinoma by predominant adenocarcinoma pattern, and relationship with presence and type of mutation and immunophenotype class.

	All patients (N=531)		Acinar/ Papillary (N=322)		Lepidic (N=45)		Micropapillary(N=17)		Mucinous (N=29)		Solid (N=118)		*p*
	No	**%**	**No**	**%**	**No**	**%**	**No**	**%**	**No**	**%**	**No**	**%**	
Age, mean (sd)	66.5 ±8.6		66.3 ±8.4		68.8 ±7.9		61.2 ±14.2		66.4 ±7.3		66.2 ±8.6		*0.04*
Follow-up, years (median, IQR)	3.55 (1.96-6.71)		3.87 (2.06-6.72)		5.91 (2.54-7.73)		6.14 (2.87-9.19)		2.44 (1.14-4.17)		2.79 (1.56-5.77)		*<0.001*
Sex													
Females	176	33.2	122	37.9	20	44.4	2	11.8	9	31.0	23	19.5	*0.001*
Males	355	66.9	200	62.1	25	55.6	15	88.2	20	69.0	95	80.5	
Smoking													
Never	81	15.3	51	15.8	12	26.7	4	23.5	3	10.3	11	9.3	
Ex	218	41.1	140	43.5	15	33.3	5	29.4	14	48.3	44	37.3	*0.006*
Ever	120	22.6	69	21.4	7	15.6	1	5.9	3	10.3	40	33.9	
missing	112	21.1	62	19.3	11	24.4	7	41.2	9	31.0	23	19.5	
Grading													
G1	54	10.2	21	6.5	29	64.4	0	0.0	4	13.8	0	0.0	
G2	299	56.3	257	79.8	13	28.9	1	5.9	16	55.2	12	10.2	*<0.001*
G3	173	32.6	43	13.4	1	2.2	16	94.1	7	24.1	106	89.8	
missing	5	0.9	1	0.3	2	4.4	0	0.0	2	6.9	0	0.0	
TNM Stage													
IA-IB	265	49.9	183	56.8	26	57.8	6	35.3	10	34.5	40	33.9	
IIA-IIB	104	19.6	58	18.0	8	17.8	4	23.5	7	24.1	27	22.9	
IIIA-IIIB	112	21.1	54	16.8	8	17.8	6	35.3	9	31.0	35	29.7	*0.013*
IV	46	8.7	24	7.5	2	4.4	1	5.9	3	10.3	16	13.6	
missing	4	0.8	3	0.9	1	2.2	0	0.0	0	0.0	0	0.0	
Vascolar invasion													
Absent	223	42.0	136	42.2	33	73.3	5	29.4	13	44.8	36	30.5	*<0.001*
Present	308	58.0	186	57.8	12	26.7	12	70.6	16	55.2	82	69.5	
Previous neoplasm													
Absent	346	65.2	212	65.8	28	62.2	15	88.2	17	58.6	74	62.7	*0.279*
Present	185	34.8	110	34.2	17	37.8	2	11.8	12	41.4	44	37.3	
Mutation													
KRAS	191	36.0	112	34.8	17	37.8	5	29.4	17	58.6	40	33.9	
EGFR	83	15.6	62	19.3	10	22.2	2	11.8	3	10.3	6	5.1	
PI3K	16	3.0	8	2.5	0	0.0	1	5.9	2	6.9	5	4.2	
Wild Type	246	46.3	179	55.6	18	40.0	10	58.8	8	27.6	67	56.8	
Fenotipo													
Alveolar	279	52.5	196	60.9	22	48.9	7	41.2	8	27.6	46	39.0	
Bronchiolar	84	15.8	38	11.8	8	17.8	1	5.9	11	37.9	26	22.0	
Mixed	84	15.8	47	14.6	7	15.6	5	29.4	7	24.1	18	15.3	*<0.001*
Null Type	16	3.0	5	1.6	1	2.2	2	11.8	0	0.0	8	6.8	
Not Evaluated	68	12.8	36	11.2	7	15.6	2	11.8	3	10.3	20	17.0	

Table 3. Characteristics of 531 patients with lung adenocarcinoma by phenotype characterization, and relationship with clinical-pathological characteristics. (* test to comparison between alveolar and bronchiolar phenotype)

	All patients (N=531)		Alveolar (N=279)		Bronchiolar (N=84)		Mixed (N=84)		Null type (N=16)		Not evaluated (N=68)		*P**
	No	**%**	**No**	**%**	**No**	**%**	**No**	**%**	**No**	**%**	**No**	**%**	***p****
Age, mean (sd)	66.5 ±8.6		65.8 ±8.7		67.7 ±7.5		68.0 ±7.8		65.4 ±9.8		66.5 ±9.7		*0.06*
Follow-up, years (median, IQR)	3.55 (1.96-6.71)		3.54 (2.09-7.16)		2.72 (1.14-5.37)		3.56 (1.72-6.61)		2.84 (1.17-5.26)		4.66 (1.95-6.97)		*0.004*
Sex													*0.08*
Females	176	33.2	105	37.6	23	27.4	28	33.3	2	12.5	18	26.5	
Males	355	66.9	174	62.4	61	72.6	56	66.7	14	87.5	50	73.5	
Smoking													*0.11*
Never	81	15.3	53	19.0	7	8.3	12	14.3	0	0.0	9	13.2	
Ex	218	41.1	113	40.5	39	46.4	27	32.1	6	37.5	33	48.5	
Ever	120	22.6	60	21.5	17	20.2	22	26.2	6	37.5	15	22.1	
missing	*112*	*21.1*	*54*	*19.0*	*21*	*25.0*	*23*	*27.4*	*4*	*25.0*	*11*	*16.2*	
TNM Stage													*0.68*
IA-IB	265	49.9	37	44.1	41	48.8	10	62.5	69	47.3	140	50.2	
IIA-IIB	104	19.6	19	22.6	18	21.4	3	18.8	30	20.6	51	18.3	
IIIA-IIIB	112	21.1	20	23.8	19	22.6	2	12.5	33	22.6	59	21.2	
IV	46	8.7	8	9.5	6	7.1	1	6.3	12	8.2	26	9.3	
missing	*4*	*0.8*	*0*	*0.0*	*0*	*0.0*	*0*	*0.0*	*2*	*1.4*	*3*	*1.1*	
Vascular invasion													*0.01*
Absent	223	42.0	128	45.6	26	30.9	28	33.3	4	25.0	37	54.4	
Present	308	58.0	151	54.1	58	69.1	56	66.7	12	75.0	31	45.6	
Previous neoplasm													*0.67*
Absent	346	65.2	181	64.9	59	70.2	54	64.3	10	62.5	42	61.8	
Present	185	34.8	98	35.1	25	29.8	30	35.7	6	37.5	26	38.2	
Mutation													*0.002*
KRAS	186	35.0	96	34.4	38	45.2	25	29.8	2	12.5	25	36.8	
EGFR	83	15.6	60	21.5	6	7.1	14	6.7	0	0.0	3	4.4	
PI3K	16	3.0	7	2.5	6	7.1	1	1.2	1	6.3	1	1.5	
Wild Type	246	46.3	116	47.2	34	13.8	44	17.9	13	5.3	39	15.9	
Adenocarcinoma patterns													
Acinar/Papillary	322	60.6	196	70.3	38	45.2	47	56.0	5	31.3	36	52.9	
Lepidic	45	8.5	22	7.9	8	9.5	7	8.3	1	6.3	7	10.3	
Micropapillary	17	3.2	7	2.5	1	1.2	5	6.0	2	12.5	2	2.9	
Mucinous	29	5.5	8	2.9	11	13.1	7	8.3	0	0.0	3	4.4	
Solid	118	22.2	46	16.5	26	31.0	18	21.4	8	50.0	20	29.4	

Table 4. Univariate and Multivariate analyses of overall survival in 531 patients with adenocarcinoma

	Univariate analysis			Multivariate analysis		
Variable	**HR**	**95% CI**	**p**	**HR**	**95% CI**	**p**
Age	1.02	1.01-1.04	0.003	1.03	1.01-1.05	<0.001
Years of surgery	1.03	0.99-1.07	0.283	0.98	0.92-1.04	0.463
Sex						
Females	1.00			1.00		
Males	1.36	1.03-1.80	0.028	1.18	0.86-1.61	0.298
Smoking						
Never	1.00			1.00		
Ex	1.30	0.87-1.93	0.197	1.20	0.77-1.88	0.417
Ever	1.35	0.86-2.08	0.175	1.37	0.84-2.24	0.210
missing	1.14	0.74-1.75	0.547	1.03	0.61-1.73	0.904
TNM Stage						
IA-IB	1.00			1.00		
IIA-IIB	2.07	1.48-2.88	<0.001	1.96	1.38-2.76	<0.001
IIIA-IIIB	2.98	2.17-4.09	<0.001	3.23	2.28-4.58	<0.001
IV	3.85	2.61-5.67	<0.001	4.38	2.89-6.63	<0.001
Vascolar invasion						
Absent	1.00			1.00		
Present	1.97	1.52-2.56	<0.001	1.51	1.13-2.03	0.006
Adenocarcinoma						
Acinar/Papillary	1.00			1.00		
Lepidic	0.99	0.63-1.55	0.963	1.24	0.77-2.00	0.384
Micropapillary	0.88	0.43-1.79	0.725	0.63	0.29-1.33	0.223
Mucinous	1.96	1.20-3.21	0.007	1.43	0.86-2.39	0.172
Solid	1.52	1.13-2.03	0.005	0.99	0.72-1.35	0.931
Mutation						
Wild Type	1.00			1.00		
KRAS	1.01	0.77-1.32	0.938	1.05	0.78-1.40	0.751
EGFR	0.80	0.55-1.17	0.256	0.99	0.65-1.51	0.957
PI3K	1.61	0.84-3.07	0.148	1.56	0.79-3.09	0.198
Phenotype						
Alveolar	1.00			1.00		
Bronchiolar	1.60	1.14-2.25	0.006	1.42	0.99-2.05	0.057
Mixed	1.45	1.04-2.02	0.028	1.38	0.98-1.95	0.067
Null Type	2.38	1.34-4.22	0.003	2.54	1.36-4.74	0.003
Not evaluated	1.01	0.68-1.51	0.955	1.02	0.67-1.56	0.913

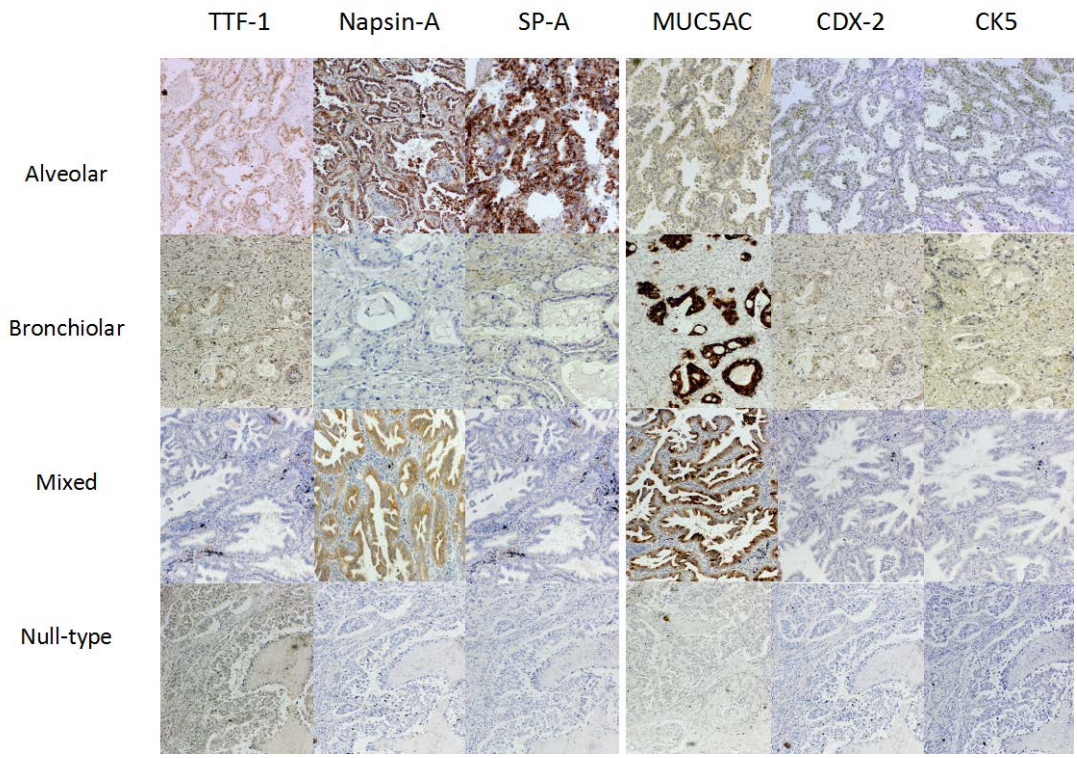
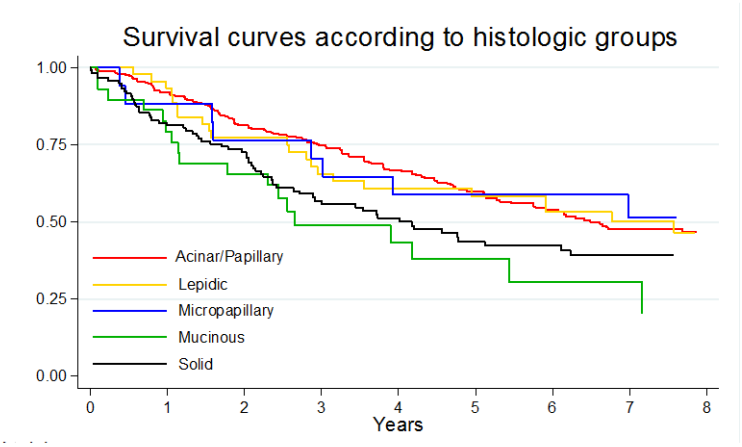
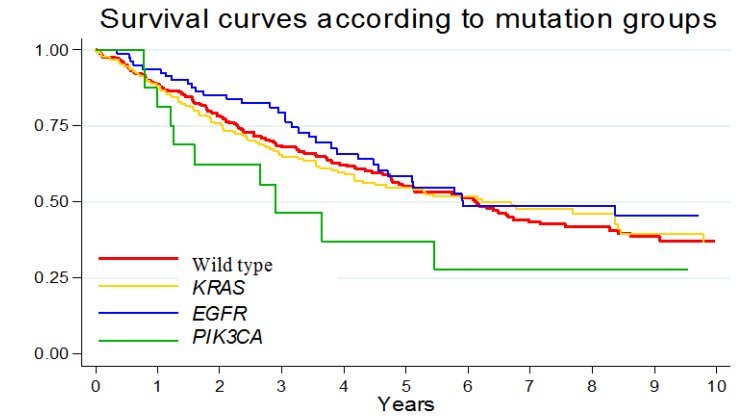


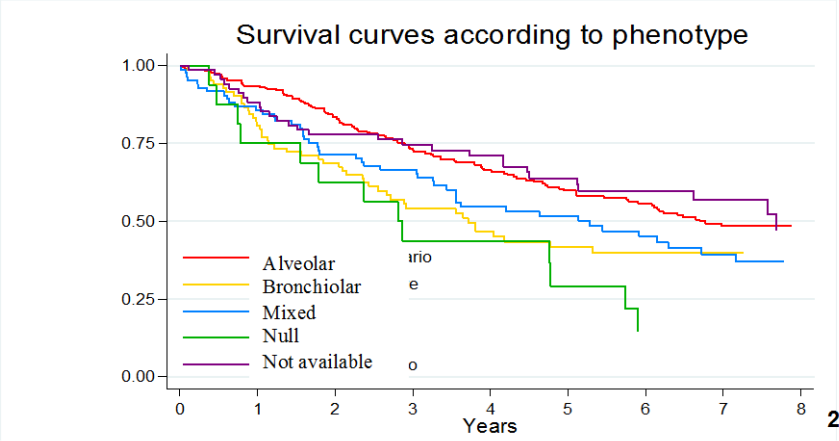
Figure 1. In this figure it has been documented the six cell-of-origin immunohistochemical markers that allowed to identify four types of lung adenocarcinoma (alveolar, bronchiolar, mixed and null types). Indeed, here it has been reported an example of this cell-of-origin profile, in which the immunohistochemistry meets the morphology. The 3 columns on the left correspond to the immunohistochemical markers of alveolar differentiation (TTF1, Napsin A and SP-A); the 3 columns on the right show the immunohistochemical markers of bronchiolar differentiation (MUC5AC, CDX2 and CK5).
Particularly, in this figure, the alveolar adenocarcinoma shows a strong and diffuse expression of TTF1, Napsin A and SP-A and was negative for all the bronchiolar markers. On the other hand, the adenocarcinoma with bronchiolar pattern does not express alveolar markers, but demonstrates a diffuse expression of MUC5AC. The mixed type of pulmonary adenocarcinoma presents positivity for a marker of alveolar diiferentiation (Napsin A) and, at the same time, also for a marker of bronchiolar differentiation (MUC5AC). The null-type adenocarcinoma does not express any of the cell-of-origin (alveolar and bronchiolar) immunohistochemical markers (original magnification: 100x).



2A



2B



2C

Figure 2. The main findings highlighted by these Kaplain-Meier curves are: 1) a better survival for ADC with acinar/papillary pattern in the curve based on histologic group (Fig. 2A), 2) similar prognosis among EGFR mutated, KRAS mutated and Wild Type patients in the curve based on mutational status (Fig.2B), 3) alveolar ADC have the best clinical course, as well as the null-type the worst one, in the curve based on phenotype (Fig.2C).

Table 1 Suppl. Hot spot gene mutations detected by Sequenom

GENE	MUTATIONS DETECTED
KRAS ex **2** cod 12	G12S/V/F/R/A/C/D
KRAS ex **2** cod 13	G13C/S/A/V/D
KRAS ex **3** cod 61	Q61L/R/P/H/E/K
BRAF ex **15** cod 594	D594G/V
BRAF ex **15** cod 600	V600E/K/M
NRAS ex **3** cod 61	Q61E/K/H/L/R/P
EGFR ex **18** cod 719	G719S/C/A/D, G719S/C/A/D
EGFR ex **19** cod 746-754	E746_T751del, E746_A750del, E746_T751del, E746_T751del, S752D, L747_E749del, L747_T750del, L747_S752del, L747_T751del, L747_S752del, P753S, A750P, T751A, T751P, T751I, S752I/F, S752_I759del, L747_Q ins, E746_T751del, I ins (combined), E746_A750del, T751A (combined), L747_E749del, A750P (combined), L747_T750del, P ins (combined), L747_S752del
EGFR ex **20** cod 768	S768I/N, A767_S768insTLA
EGFR ex **20** cod 790	T790M
EGFR ex **20** cod 770-771	V769_D770insASV, V769_D770insCV, D770_N771>AGG/V769_D770insASV/V769_D770insASV, D770_N771insG,
EGFR ex **21** cod 858	L858R/M
EGFR ex **21** cod 861	L861Q
PIK3CA ex **9** cod 542	E542Q/K
PIK3CA ex **9** cod 545	E545Q/K
PIK3CA ex **20** cod 1047	H1047Y/R/L