Molecular characteristics of ERCC1 negative vs. ERCC1 positive tumors in resected NSCLC

Luc Friboulet\textsuperscript{1,2,3}, Daniel Barrios-Gonzales\textsuperscript{2,3,4}, Frédéric Commo\textsuperscript{1,2,3}, Ken André Olaussen\textsuperscript{1,2,3}, Stephan Vagner\textsuperscript{1,2,3}, Julien Adam\textsuperscript{1}, Aïcha Goubar\textsuperscript{1,2,3}, Nicolas Dorvault\textsuperscript{1,2,3}, Vladimir Lazar\textsuperscript{5}, Bastien Job\textsuperscript{5}, Benjamin Besse\textsuperscript{2,3,4}, Pierre Validire\textsuperscript{6}, Philippe Girard\textsuperscript{6}, Ludovic Lacroix\textsuperscript{1,7}, Johanna Hasmats\textsuperscript{8}, Fabienne Dufour\textsuperscript{4}, Fabrice André\textsuperscript{1,2,3,4}, Jean-Charles Soria\textsuperscript{1,2,3,4}

\textsuperscript{1}INSERM, U981, Villejuif, France; \textsuperscript{2}Institut de cancérologie Gustave Roussy, Villejuif, France; \textsuperscript{3}Université Paris-Sud XI, Kremlin-Bicêtre, France; \textsuperscript{4}Department of medicine, Institut Gustave Roussy, Villejuif, France; \textsuperscript{5}Unité de Génomique Fonctionnelle et Bioinformatique, Institut Gustave Roussy, Villejuif, France; \textsuperscript{6}Département Thoracique, Institut Mutualiste Montsouris, Paris, France; \textsuperscript{7}Laboratoire de Recherche Translationnelle, Institut Gustave-Roussy, Villejuif, France; \textsuperscript{8}Science for Life Laboratory, School of Biotechnology, Division of Gene Technology, Royal Institute of Technology, Stockholm, Sweden.

On Behalf of Chemores Consortium: FP6

Corresponding author: Jean-Charles Soria (MD/PhD); INSERM U981; Institut Gustave Roussy; 114, rue Edouard Vaillant F-94805 Villejuif, France; Phone: 33-1-42-11-43-39; Fax: 33-1-42-11-52-17; E-mail: soria@igr.fr

Running title: ERCC1 negative vs. positive tumors in resected NSCLC

Key-words: ERCC1, genomic instability, NSCLC, \textit{EGFR} mutations, qRT-PCR/IHC.
Conflict of interest: All authors certify that they have no conflict of interest. All authors contributed to design, execution or analysis of experiments. FC has accomplished most statistical analysis. LF has written most parts of the manuscript. JCS has designed and supervised the study.

Word count (excluding references), translational relevance 148 / 150; abstract: 245 / 250; text: 4489 / 5000; number of figures: 4 and table: 1
Statement of translational relevance

Adjuvant chemotherapy has a significant, but limited effect in prolonging survival. The use of biomarkers to select patients and optimize treatment options is essential. ERCC1 is a promising biomarker but the ID-card of ERCC1 positive vs. negative tumors is unclear.

To better characterize ERCC1 positive vs. negative tumors, we performed a global molecular analysis on 91 resected NSCLC samples. The present study addresses some of the current debates in the literature, notably it performs a comparison between IHC and qRT-PCR to evaluate ERCC1 expression level, and estimates the link between cancer-gene mutations (i.e. EGFR) and ERCC1 status.

The differences between ERCC1 negative and positive tumors might help to apprehend the underlying mechanisms of chemoresistance. Our data suggest that ERCC1 status could be associated with higher genomic instability with a specific pattern of DNA-damage-response gene expression. Furthermore miRNA expression may explain discrepancies between IHC and mRNA measurements of ERCC1.
Abstract

PURPOSE: ERCC1 is a protein involved in repair of DNA platinum adducts and stalled DNA replication forks. We and others have previously demonstrated the influence of ERCC1 expression upon survival rates and benefit of cisplatin-based chemotherapy in patients with resected non-small cell lung cancer (NSCLC). However little is known about the molecular characteristics of ERCC1 positive and negative tumors.

EXPERIMENTAL DESIGN: We took advantage of a cohort of 91 patients with resected non-small cell lung cancer for which we had matched frozen and paraffin-embedded samples to explore the comparative molecular portraits of ERCC1 positive and negative tumors of NSCLC. We performed a global molecular analysis including assessment of ERCC1 expression levels using both immunohistochemistry (IHC) and quantitative RT-PCR, genomic instability (CGH), global gene and miRNA expression and sequencing of selected key genes involved in lung carcinogenesis.

RESULTS: ERCC1 protein and mRNA expression were significantly correlated. However, we observed several cases with clear discrepancies. We noted that ERCC1 negative-tumors had a higher rate of genomic abnormalities vs. ERCC1 positive-tumors. ERCC1 positive-tumors seemed to share a common DNA-damage response phenotype with the overexpression of 7 genes linked to DNA-damage response. The miRNA expression analysis identified miR-375 as significantly under-expressed in ERCC1 positive-tumors.

CONCLUSIONS: Our data demonstrate inconsistencies in ERCC1 expression between IHC and qRT-PCR readouts. Further, ERCC1 status is not linked to specific mutational patterns or frequencies. Finally, ERCC1-negative tumors have a high rate of genomic aberrations that could consequently influence prognosis in patients with resected NSCLC.
Introduction

Lung cancer is the leading cause of cancer-related deaths worldwide. It accounts for 13% of all cancers and 18% of the deaths in 2008, and non-small cell lung cancer (NSCLC) accounts for approximately 85% of all lung cancer cases (1). Cisplatin (and its analogs) remain the mainstay of chemotherapy in many solid tumors. However, platinum-based chemotherapy, in both local and metastatic disease, has only modest effect in prolonging survival and is associated with significant adverse effects (2-4). A way to improve clinical outcome is to identify biomarkers in tumor cells whose expression pattern is able to predict clinical benefit from cisplatin-based chemotherapy. Such a strategy would allow the selection of patients with improved chance of response and survival. Various predictive biomarkers are currently being investigated for their ability to guide treatment decisions. Successful application of oriented targeted therapies such as Epidermal Growth Factor Receptor (EGFR) inhibition further supports the importance of understanding the molecular heterogeneity in lung cancer (5).

Platinum based compounds induce their cytotoxic effects by covalent binding to DNA, forming intra- and inter-strand platinum-DNA adducts. Without efficient DNA repair, these adducts interfere with DNA transcription and replication, eventually inducing cell death (6). Therefore, DNA damage repair pathways, in addition to their implications in genomic instability and tumorigenesis, are highly relevant to predict platinum resistance (7). There are at least 6 main DNA repair pathways known to correct DNA damage: Direct Repair (DR), Base Excision Repair (BER), Nucleotide Excision Repair (NER), Mismatch Repair (MMR), Homologous Recombination Repair (HRR), Non-Homologous End Joining (NHEJ). The mechanism of repair mobilized depends on the type of DNA-damage and the cell cycle phase.

The excision repair cross-complementation group 1 (ERCC1) gene encodes a key enzyme in the NER pathway, which recognizes and removes platinum-DNA adducts (8). The ERCC1/XPF heterodimer is responsible for incising DNA at the 5’ site of the lesion during
NER. ERCC1/XPF is also involved in the repair of platinum-induced DNA interstrand crosslinks (ICL) to allow the release of stalled replication forks.

We and others have demonstrated that ERCC1 expression is both a prognostic and predictive marker in NSCLC (9, 10). ERCC1-negative expression in tumors is predictive of survival benefit from cisplatin-based adjuvant chemotherapy, whereas in the absence of therapy patients with ERCC1-negative tumors have a worst prognosis.

ERCC1 value has also been explored in gastric cancer (11), esophageal cancer (4), ovarian cancer (12), head and neck cancer (13), bladder cancer (2) and colorectal cancer (14). In these tumors, a high level of ERCC1 mRNA or protein expression correlated with poor overall survival and resistance to platinum-based chemotherapeutic treatments. Nevertheless, some studies have not been able to correlate ERCC1 expression level with survival (15, 16). Both IHC and qRT-PCR are usually used to assess ERCC1 expression. A correlation between ERCC1 mRNA (qRT-PCR) and protein expression (IHC) has been observed in one cohort (17), but not in others (18). Thus, the choice of one technique to the other could explain the discrepancies between different studies.

Several studies have also investigated whether or not the presence of EGFR mutation was related to the prognosis of resected NSCLC (3, 19). Interestingly, EGFR activating mutation were associated with low ERCC1 expression both at the mRNA (20) and protein level (19). This is in contrast with previous studies that suggested that ERCC1 expression was upregulated by EGFR pathway activation (21, 22).

Little is known about the specific molecular differences between ERCC1 positive vs. negative tumors. Therefore, we aimed at obtaining a better picture of the molecular background associated with a specific ERCC1 status. To identify molecular differences between ERCC1 positive and negative NSCLC tumors we performed a global molecular analysis on a cohort of 91 resected NSCLC patients for which we had matched frozen and
paraffin-embedded tissues. We analyzed the correlation between ERCC1 IHC status and (a) ERCC1 mRNA expression level (qRT-PCR), (b) genomic instability (CGH), (c) global gene expression (transcriptome analysis microarrays), (d) miRNA expression (miRNA microarrays) and (e) sequencing of important genes involved in lung carcinogenesis (TP53; KRAS; EGFR; PIK3CA; BRAF; ERBB2; KDR; AKT1 and ERCC1).
Materials and methods:

Patients and samples

NSCLC patients underwent surgical resection in a single institution (Institut Mutualiste Montsouris in Paris, France) between 2002 and 2006. The key inclusion criterion was the availability of matched formalin-fixed paraffin-embedded (FFPE) and frozen samples. The present cohort belongs to the CHEMORES initiative (Chemotherapy resistance consortium), which is an EU funded (FP6) research collaboration involving 19 universities, organizations for cancer research, and research-oriented biotechnology companies in 8 European countries.

ERCC1 immunohistochemistry

FFPE tissue sections were stained with mouse monoclonal antibody against ERCC1 (ab-2, clone 8F1, ref: MS-671-P, Thermo scientific, Essex, UK) as previously described (9). Microscopic evaluation of the ERCC1 staining was performed independently by three observers (D.B., K.A.O., J.A.) in a blinded fashion. The staining intensity was graded on a scale of 0 to 3 based on the intensity of internal control cells (endothelium, bronchial epithelium or fibroblast cells close to the tumor) that was systematically assigned an intensity of 2. The percentage of positive nuclei was estimated for each specimen, and a proportional score was assigned (0 if 0%, 0.1 if 1-9%, 0.5 if 10-49% and 1 if 50% or more). The proportional score was multiplied by the staining intensity to obtain a final semi-quantitative H Score, as previously described (9).

RNA extraction and quantitative reverse transcriptase PCR (qRT-PCR)

No macrodissection or microdissection of the tissues before mRNA extraction was performed. The lysis of each NSCLC frozen tissue sample was obtained using a polytron Homogeniser.
The RNA extraction was performed with TRIzol® Reagent Protocol (Invitrogen, Carlsbad, CA, USA). Total RNA was quantified and qualified with Nanodrop ND-1000 spectrometer and Bioanalyser-2100 (Agilent, Santa Clara, CA, USA). The primers and probe sequences used for the quantitative PCR for ERCC1 were the following: Forward primer: 5’-GGGAATTTGGCGACGTAATTC-3’; Reverse primer 5’-GCGGAGGCTGAGGAACAG-3’, probe 6FAM 5’-CACAGGTGCTCTGGCCCAGCACATATAM-3’. For PPIA: 5’-GTCAACCCCACCCTTGTCTTTCT-3’, 5’-CTGCTGTCTTTTGGGACCTTGT-3’, 6FAM 5’-AGCTCAAGGAGACCGGCCCAC-TAM-3’, for the forward, reverse primer, and probe, respectively. The ERCC1 primers recognize exon 5 of the ERCC1 gene common to all known ERCC1 isoforms. The relative expression of ERCC1 mRNA was determined using the Ct value and the 2^{-ΔΔCt} method. The data are presented as the fold-change in gene expression normalized to an endogenous reference gene (PPIA) and relative to a human reference cDNA.

CGH

The DNA copy number profiling was carried out by dual-color competitive hybridization on 244K Whole Human Genome (G4411B) microarrays (Agilent Technologies, Santa Clara, CA, USA). In each case, the normal tissue sample was used as the reference to its corresponding tumor sample. Tumor aberration rates were defined as the ratio between the total length of aberrations in a given tumor sample versus the full size of the human genome as defined by the UCSC genome build hg18. See Supplemental methods for Experimental procedures (23).

Gene expression microarrays

Gene expression profiling was performed using dual-color 244K Human Exon Array (design-custom) 24891 from Agilent. Each normal sample was used as the reference to its
corresponding tumor sample. RNA extractions, Cy3 (normal samples) and Cy5 (tumor samples) labeling, and hybridizations were carried out according to Agilent’s instructions. Slides were scanned with an Agilent MicroArray Scanner G2505B (Agilent Technologies Inc, Santa Clara, CA), and Feature Extraction normalized data were imported into R software. See Supplemental methods for analysis details.

**miRNA microarrays**

The miRNA expression profiling of tissue samples was performed on Agilent Human microRNA microarray v3.0. According to Agilent’s recommendations, normal and tumor samples were hybridized separately. Briefly, isolated total RNA were dephosphorylated with Calf Intestinal Phosphatase, labeled with pCp-Cy3 and T4RNA Ligase, and hybridized for 20h at 55°C. Slides were washed and scanned with an Agilent MicroArray Scanner G2505B. Feature Extraction data were imported into R software, then quantiles-normalized and log2-transformed before statistical analysis.

**Gene mutations analysis**

A search for mutations by sequencing of 9 target genes was performed internally by the translational research laboratory at Institut Gustave Roussy, (Villejuif France) except for p53 status which was analyzed by the Royal Institute of Technology (Stockholm, Sweden). We analyzed full coding sequences of exons including oncogenic mutational hotspots in NSCLC corresponding to: *TP53* (NM_000546.4) exons 5-8; *KRAS* (NM_004448.2) exons 2,3; *EGFR* (NM_005228.3) exons 18-21; *PIK3CA* (NM_006218.2) exons 10, 21; *BRAF* (NM_004333.4) exon 15; *ERBB2* (NM_004448.2) exons 18,20-24; *KDR* (NM_002253.1) exons 2, 26,27,30; *AKT1* (NM_005163.2) exon 4 and *ERCC1* (NM_202001.1) all exons. See Supplemental methods for Experimental procedures.
**Statistical analysis**

All statistical analysis were performed with R software. A SAM procedure (R software, samr package) was carried out to define genes or miRNA differentially expressed in tumors regarding their ERCC1-IHC status. A q.value(%) ≤ 5 was considered as significant. For CGH, frequencies of gain, or loss, were tested for their association with the ERCC1-IHC status using a Pearson’s Chi² test using a Monte-Carlo approximation on 2000 replicates. P-values were then corrected with the Benjamini & Hochberg method (24). A p-value(%) ≤ 5 was considered as significant. CGH aberration rates and qRT-PCR values were considered as continuous values and were compared regarding ERCC1-IHC status with a Wilcoxon test. The Kaplan-Meier method was used to estimate the overall survival curves according to CGH aberration rate. For this, patients were assigned into low or high aberration rate group by the Kmeans clustering method. A log-rank test was used to compare rates between groups.
Results

Clinicopathological characteristics

Among 121 initial tumors stained for ERCC1, 14 were excluded from further analysis because of invalid staining of internal control cells during IHC staining and 16 were excluded because they contained less than 50% of tumor cells. All molecular and statistical analyses were thus performed on the 91 remaining samples. The clinicopathological characteristics of the 91 included patients are shown in Table 1. The median age was 63 years. Patients were more frequently men in both ERCC1 positive and negative groups (70-73%). Pathological stages were stage I in 46 cases (52%), stage II in 18 cases (20%) and stage III in 25 cases (28%). In the ERCC1 negative group, tumors were predominantly adenocarcinomas (61%) while in the ERCC1 positive group, there was a majority of squamous cell carcinomas (SCCs) (60%, P = 0.001) as previously reported (9). Almost half of the patients received adjuvant chemotherapy and 88% did not receive adjuvant radiotherapy.

Patient classification according to IHC assessment of ERCC1 expression

The median percentage of cells with positively stained nuclei was 60% (range 0 to 90%) whereas the median value of intensity was 2. Tumors with an H-score exceeding 1.5 were considered ERCC1 positive as previously reported (9). Among the 91 tumors, 50 were classified as positive and 41 as negative. Figure 1A display an example of ERCC1 nuclei staining with a positive staining intensity of 3. In this small cohort, ERCC1 protein expression level had no prognostic or predictive value (data not shown). ERCC1 status reported in the rest of this manuscript was defined by IHC unless otherwise specified.

Correlation between ERCC1 mRNA (qRT-PCR) and protein expression levels (IHC)
We compared ERCC1 H-scores determined by IHC to the relative amount of all known ERCC1 mRNA transcripts determined by qRT-PCR. ERCC1 IHC status was significantly associated with the relative amount of ERCC1 mRNA transcripts (Figure 1B). The median of ERCC1 $2^{-\Delta\Delta C_t}$ was 0.75 in the ERCC1 IHC-negative group compared to 1.06 in the ERCC1 positive group ($p = 0.003$). A logistic regression model confirmed the correlation between the relative expression of ERCC1 mRNA and the probability of ERCC1 IHC-positive status ($p = 0.004$, Figure 1C). The median mRNA expression (0.92) was chosen as a cut-off point to identify patients as “low ERCC1” or “high ERCC1” (as previously described) (25). Figure 1D illustrates the spectrum of ERCC1 mRNA expression levels, centered on the global median, according to ERCC1 IHC status with a significant difference between both groups (khi-square, $p<10^{-3}$). Of note, Figure 1D suggests that this correlation is far from being perfect, since 33% of the samples were ERCC1 low by qRT-PCR in ERCC1 IHC-positive tumors. In ERCC1 IHC-negative tumors 32% of the samples were high for ERCC1 by qRT-PCR. We therefore investigated whether these disparities could be explained by the presence of ERCC1-targeted miRNA differentially expressed between ERCC1 positive and negative tumors.

**Differential micro-RNA expression according to ERCC1 status**

We performed a global miRNA expression analysis. The best patient classification regarding ERCC1-IHC status was obtained by considering the first 29 miRNA when ordered by q-values (Fisher exact test, $p = 3.10^{-3}$ on 29 miRNA, Figure 2A and table S1). When selecting q-values lower than $10^{-3}$ for each individual miRNA, 4 were expressed differentially between the two groups. Among these, hsa-miR-375 was the most differentially expressed with a fold change of -3.3 in the positive group compared to the negative one (Figure 2B). None of these 4 miRNA has a predicted interaction with ERCC1 mRNA according to TargetScan, release 5.1 (http://www.targetscan.org). With other target prediction program (miRGen:
we identified, among the genes predicted to be targeted by miR-375, 9 genes linked to DNA repair: *TP53, USP1, APEX1, TYMS, MLH3, XRCC6BP1, ERCC3, NTHL1* and *PARP4*.

However, when searching exclusively for the presence of ERCC1-targeting miRNA that could explain disparities between ERCC1 protein and mRNA expression levels, we identified hsa-miR-409-3p. Indeed, the expression of miR-409-3p increased in parallel with ERCC1 mRNA expression specifically in ERCC1 negative-tumors (Figure 2C), whereas, it had a stable expression in ERCC1 positive-tumors. mi-RNA can decrease the protein expression, by adapting its own expression to mRNA expression of the target gene. Here, miR-409-3p could maintain low ERCC1 protein expression in ERCC1 negative-tumors even when ERCC1 mRNA expression increased. This phenomenon can contribute to the observed discrepancies between IHC and qRT-PCR results.

**ERCC1 expression influences genomic aberration rates (CGH)**

The DNA copy number profiling showed some weak differences in gain/loss frequencies between the 2 ERCC1 IHC statuses. Precisely, we observed higher frequencies of gains in 8q and 14q regions in ERCC1 IHC-positive tumors. But both appeared non significant after p-value correction (Figure S1).

However, the median of aberration rates (defined as the sum of lengths of aberrations related to the size of explored genome, for each tumor sample (categorized either as a relative gain or loss of DNA copy number)) (Figure 3A) was 43.2% in the ERCC1 negative group (range from 0.02 to 78.5) compared to 24.2% in the ERCC1 positive group (range 0.05 – 72.4) (Figure 3B). This difference in rates, regarding ERCC1 IHC status, appeared significant (*p = 0.004*), and suggests a higher genetic instability among ERCC1 negative tumor samples as compared to positive ones. Interestingly, the analysis on overall survival according to the
overall rate of genomic aberrations in the tumors (CGH) suggested that genomic instability was not associated with survival (p = 0.47), but high aberrant tumors clearly tended to have a worst prognosis (Figure 3C).

**ERCC1 protein expression is not associated with a specific pattern or frequency of mutations in relevant cancer-related genes**

We then examined if ERCC1 IHC status reflected genomic instability in terms of number of mutations. We analyzed the number of mutations by direct sequencing of mutational hotspots in 9 key genes in NSCLC (EGFR, KRAS, TP53, PI3CA, BRAF, AKT1, KDR, ERBB2 and ERCC1). Several mutations were detected, but no significant differences were observed between ERCC1 positive and negative groups (Figure 3D). Only one mutation in ERCC1 exon 8 was found. Concerning EGFR, 6 activating mutations were identified and for TP53 we found mutations in 27% of patients. The details of mutations detected are presented in table S2 (26).

**Differential gene expression between ERCC1 positive and negative tumors**

We carried out a global transcriptome analysis to evaluate whether there were any significant differences in gene expression between ERCC1 positive and ERCC1 negative tumors. We found 149 genes that were differentially expressed between the two groups (Figure 4A and Table S3). None of these genes have previously been described to belong to the NER pathway or any of the other DNA repair pathways. However, we found 7 genes strongly associated with DNA-damage response (DDR) and all were increased in the ERCC1 positive group. These DDR-related genes were CCNA1, KLF4, MYC, RNF168, S100A2, SNAI2 and TP53AIP1 (GO: 0006974) (Figure 4B). Three genes overexpressed in ERCC1 positive tumors are linked to the EGFR signaling pathway ITGA6, KRT16, KRT6B (GO:0007173).
Discussion

One of the major challenges in current cancer chemotherapeutic treatment is the identification of effective biomarkers for drug response prediction (27). Although many studies suggested that ERCC1 is a potential prognostic and predictive marker, there is no consensus on the optimal method for ERCC1 evaluation (IHC vs. qRT-PCR) and little is known regarding the molecular basis related to the intrinsic prognostic value of ERCC1. Here we performed a “system biology” approach including CGH arrays, gene expression arrays, miRNA microarrays, and gene sequencing analysis in 91 resected NSCLC to better apprehend how ERCC1 IHC-positive tumors might be different from their negative counterparts.

In several malignancies ERCC1 mRNA expression has shown to be related with clinical outcome and response to DNA damaging chemotherapy (28). Metzger et al reported increased ERCC1 mRNA expression as an indicator for non response to neoadjuvant cisplatin-based chemotherapy (29). However, they did not study the correlation between ERCC1 expression at the protein and transcriptional level. We found a certain degree of association between ERCC1 mRNA levels and ERCC1 protein expression observed by IHC. This association was somehow imperfect and the overlap between the populations was considerable. In line with our results, Doll et al. have described a similar degree of correlation between ERCC1 mRNA and protein expression, with clear areas of discrepancies (17). Another study was unable to correlate ERCC1 mRNA and protein expression (18). These disparities can be explained by technical reasons. Indeed, the IHC approach takes into account the percentage and intensity of protein expression directly in tumor cells. On the contrary, the qRT-PCR depends on the quality of the sample (RNA is easily altered) and the percentage of tumor cells in the sample varies from one sample to another. It should be noted that no macrodissection or microdissection of the tissues was performed in our samples before mRNA extraction. Biological reasons can also contribute to the observed discrepancies.
Considering the importance of the post-transcriptional regulatory mechanisms of ERCC1 expression, the mRNA expression level could be very distant from the protein level. Indeed, we observed that an increase in ERCC1 mRNA expression was not always associated with an increase in protein expression. Powley et al. have proposed that levels of ERCC1 proteins are highly controlled at the translational level (30). Here, we identified the putative ERCC1-targeting miR-409-3p that behaved differently in ERCC1 positive vs. negative-tumors (see Figure 2C).

Concerning the IHC evaluation of ERCC1 expression, the choice of the optimal antibody has been largely discussed in the literature and no consensus has yet been established. ERCC1 antibodies appear to be particularly sensitive to incubation conditions and the technical staining procedures have to be optimized for each antibody (31). Furthermore, the ERCC1 epitopes recognized by the different ERCC1 antibodies are unknown. ERCC1 IHC and qRT-PCR discrepancies could therefore be also related to the presence of different ERCC1 isoforms (32). The antibodies used may be specific to one or several ERCC1 isoforms while qRT-PCR primers designed in this study recognized all ERCC1 isoforms.

We further established that ERCC1 negative tumors had a higher rate of genomic aberrations, thus, sustaining a prior hypothesis (9, 28) suggesting that ERCC1 would influence genomic stability, with lower number of genomic alterations in DNA in ERCC1-positive tumors. Such a correlation between ERCC1 loss and genomic instability has previously been suggested by a micronuclei study on mice (33). The authors proposed that the increased level of unrepaired lesions and double strand breaks in ERCC1 deficient mouse fibroblast leads to an increase in mutation frequency and genome instability. As a result of their control function it is obvious that DNA repair proteins are essential for maintaining the genome integrity (7). However, prior to our analysis, no data were available concerning a
direct association between genomic instability and ERCC1 IHC status in NSCLC tumors. Our results, however, should allow a better understanding of the intrinsic prognosis value of ERCC1 protein expression in resected NSCLC.

We also investigated if the presence of mutations was related to ERCC1 expression. TP53 mutations were the most frequently observed (27% of patients). This elevated frequency of TP53 mutations is common in NSCLC but the influence of TP53 mutations on survival and drug resistance is largely controversial (26, 34-38). Concerning the EGFR gene, many reports have described that EGFR inhibitors (erlotinib, gefitinib…) increase survival in NSCLC patients harboring activating EGFR mutations (39). We thus searched for EGFR mutations at a mutational "hotspot" within the kinase domain of the EGFR gene (exons 18-21) (40). We did not observe increased number of mutations (for EGFR and the other genes) in ERCC1 negative tumors. This is in contrast with previous studies which directly linked EGFR mutations with low ERCC1 expression (19, 20). The low number of EGFR mutations found in this study can be explained by the low percentage of never-smokers in this cohort (3%) (Table1). EGFR mutations have indeed been linked to the smoking status of patients (41).

Altogether, our results suggest that ERCC1 influences the rate of gains/losses of large genome segments rather than the rate of single nucleotide modifications. The underlying molecular mechanisms remain unclear, but ERCC1 is a general structure-specific endonuclease implicated in several important processes assuring DNA integrity, such as single-strand annealing, Fanconi anemia (FA)-dependent interstrand crosslink (ICL)-repair, recombination processes and telomere protection (42).

Our study also showed that ERCC1 positive tumors do not have a strong distinct pattern of gene expression in key pathways such as HER signaling, PI3K-PTEN-AKT-mTOR signaling or the RAF-RAS-MEK pathway. However, we identified 24 genes linked to cell
death (GO:0008219) and 3 genes overexpressed in ERCC1 positive tumors that were linked to EGFR signaling pathway ITGA6, KRT16, KRT6B (GO:0007173). Interestingly, ERCC1 positive tumors seemed to share a common DNA-damage response phenotype leading to the overexpression of 7 genes strongly linked to DNA-damage response (GO: 0006974).

We furthermore identified a 29 miRNA signature in which there was one remarkable miRNA, miR-375, which was downregulated in ERCC1 positive patients. Interestingly, miR-375 was shown to be downregulated in gastric and hepatocellular cancer (HCC) tissues compared to adjacent normal tissues (43, 44). Furthermore, miR-375 inhibited proliferation and invasion of HCC cells (44). Recent reports have suggested that miR-375 may function as a tumor suppressor by regulating cell proliferation through targeting of the Janus kinase 2 (JAK2) oncogene (43). Moreover, it has been suggested that PDK-1 (3′-phosphoinositide-dependent protein kinase-1), a key component in the Akt signaling pathway that regulates cancer cell survival and proliferation, is another direct target of miR-375 (45). Altogether, we cannot exclude that miR-375 downregulation in ERCC1 positive tumors induces a proliferative advantage in these tumors. With a target prediction program (miRGen) we identified, among the genes predicted to be targeted by miR-375, 9 genes linked to DNA repair: TP53, USP1, APEX1, TYMS, MLH3, XRCC6BP1, ERCC3, NTHL1 and PARP4. Thus, downregulation of this miRNA could also contribute to the increased DNA-damage response phenotype in ERCC1 positive tumors.

In summary, this study aimed to establish the ID-card of ERCC1 negative vs. ERCC1 positive tumors in resected NSCLC. Our data suggests that miRNA may play a role in the regulation of ERCC1 protein expression. Remarkably, ERCC1 protein expression was associated with genomic instability in terms of increased genomic aberration rates. Genomic instability has been proposed to be both a cause and a consequence of carcinogenesis.
Therefore, an association between ERCC1 and genomic instability can provide new insight into the pathogenesis of NSCLC tumors, particularly in our effort to explain the intrinsic prognostic impact of ERCC1 in untreated patients. ERCC1 status alone did not have a significant prognostic or predictive value in this small cohort. There is increasing evidence that combination of biomarkers representing multiple pathways of DNA repair increases the predictive power of these biomarkers. DNA repair is a multi-player multi-dimensional process, which evaluation should not be restricted to ERCC1. In fact a multi-parametric evaluation of different key players of DNA repair (NER/ERCC1, BER/PARP, HR/BRCA, etc) should be encouraged along with functional tests of DNA repair capacity or global read outs of DNA repair dysfunctionality (i.e. rate of aberrant genome).

**Acknowledgments:** This work was supported by grants from CHEMORES EU FP6 and Institut National du Cancer INCa “projet libre”.
Figures legends

Figure 1. Positive association between ERCC1 protein and mRNA expression level. A) Immunohistochemical assessment of ERCC1 expression. An ERCC1-positive lung tumour shows a staining intensity of 3, the highest level on a scale of 0 to 3. B) Boxplot repartition of ERCC1 mRNA expression according to ERCC1 IHC status. C) Logistic regression model of ERCC1 IHC status according to mRNA expression level (qRT-PCR). D) Histogram depicting the distribution of mRNA expression level in patients according to their ERCC1 IHC status (qRT-PCR values were centered to the median 0.92).

Figure 2. miRNA expression analysis in ERCC1 positive vs. negative tumors. A) Heatmap and 2D hierarchical clustering (Ward distance) for miRNA analysis. The first 29 differential miRNA were used to obtain an optimized patient clustering signature. B) Boxplots repartition for the four significant miRNA expressions according to ERCC1 IHC status (q-value < 10^-3). Only miR-375 had an important fold change (-3.3). C) Boxplot repartition for miR-409-3p expression according to ERCC1 IHC status and ERCC1 qRT-PCR expression level. Red dots represent miR-409-3p expression level with trendlines (blue curves).

Figure 3. ERCC1 expression influences genomic aberration rates. A) Example of a Chromosome 9’s genomic profile given by CGH, the genomic aberrations are shown in red (amplifications) and green (losses), and the thresholds of significance are shown by the dotted yellow line. To calculate the percentage of total aberrant genome per sample we divided total aberrant megabases by total analyzed genome size in each patient. B) Boxplot repartition of ERCC1 genomic aberration rates according to ERCC1 IHC status. C) Overall survival Kaplan Meier Curve according to CGH aberration rate in tumors. D) ERCC1 status and mutations
analysis on 9 cancer-genes: TP53 (NM_000546.4) exons 5-8; KRAS (NM_004448.2) exons 2,3; EGFR (NM_005228.3) exons 18-21; PIK3CA (NM_006218.2) exons 10, 21; BRAF (NM_004333.4) exon 15; ERBB2 (NM_004448.2) exons 18,20-24; KDR (NM_002253.1) exons 2, 26,27,30; AKT1 (NM_005163.2) exon 4 and ERCC1 (NM_202001.1) all exons.

**Figure 4. Differences in gene expression between ERCC1 positive and negative tumors.**
A) Heatmap and 2D hierarchical clustering (Ward distance) for transcriptome analysis. B) Boxplots repartition of mRNA expression according to ERCC1 IHC status for 7 genes associated with DNA-damage response.

**Supplementary figure and tables.**

**Figure S1.** Frequencies of gain, or loss, according to ERCC1-IHC status.

**Table S1.** List of Top 29 differentially expressed miRNA between the two groups ERCC1 positive / negative used for hierarchical clustering of patients.

**Table S2.** Mutational details.

**Table S3.** List of Top 149 significantly differentially expressed genes between the two groups ERCC1 positive / negative.
References


42. Kirschner K, Melton DW. Multiple roles of the ERCC1-XPF endonuclease in DNA repair and resistance to anticancer drugs. Anticancer research;30: 3223-32.
Table 1. Patients’ clinicopathological characteristics.

<table>
<thead>
<tr>
<th>Category</th>
<th>All Patients n=91</th>
<th>Patients with ERCC1 IHC Negative Tumors n=41</th>
<th>Patients with ERCC1 IHC Positive Tumors n=50</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>26 (28.6)</td>
<td>11 (26.8)</td>
<td>15 (30.0)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>65 (71.4)</td>
<td>30 (73.2)</td>
<td>35 (70.0)</td>
<td></td>
</tr>
<tr>
<td>Age**</td>
<td></td>
<td></td>
<td></td>
<td>0.54</td>
</tr>
<tr>
<td>Median (range)</td>
<td>62.9 (40.9 - 84.7)</td>
<td>61.4 (45.3 - 84.7)</td>
<td>63.3 (40.9 - 81.8)</td>
<td></td>
</tr>
<tr>
<td>Stage</td>
<td></td>
<td></td>
<td></td>
<td>0.57</td>
</tr>
<tr>
<td>Stage I</td>
<td>46 (51.7)</td>
<td>19 (46.3)</td>
<td>27 (56.3)</td>
<td></td>
</tr>
<tr>
<td>Stage II</td>
<td>18 (20.2)</td>
<td>10 (24.4)</td>
<td>8 (16.7)</td>
<td></td>
</tr>
<tr>
<td>Stage III</td>
<td>25 (28.1)</td>
<td>12 (29.3)</td>
<td>13 (27.1)</td>
<td></td>
</tr>
<tr>
<td>TNM Status</td>
<td></td>
<td></td>
<td></td>
<td>0.62</td>
</tr>
<tr>
<td>T: 1</td>
<td>11 (12.2)</td>
<td>3 (7.3)</td>
<td>8 (16.3)</td>
<td></td>
</tr>
<tr>
<td>T: 2</td>
<td>62 (68.9)</td>
<td>30 (73.2)</td>
<td>32 (65.3)</td>
<td></td>
</tr>
<tr>
<td>T: 3</td>
<td>10 (11.1)</td>
<td>5 (12.2)</td>
<td>5 (10.2)</td>
<td></td>
</tr>
<tr>
<td>T: 4</td>
<td>7 (7.8)</td>
<td>3 (7.3)</td>
<td>4 (8.2)</td>
<td></td>
</tr>
<tr>
<td>N: 0</td>
<td>50 (55.2)</td>
<td>23 (56.1)</td>
<td>27 (56.3)</td>
<td>0.70</td>
</tr>
<tr>
<td>N: 1</td>
<td>21 (23.6)</td>
<td>11 (26.8)</td>
<td>10 (20.8)</td>
<td></td>
</tr>
<tr>
<td>N: 2</td>
<td>18 (20.2)</td>
<td>7 (17.1)</td>
<td>11 (22.9)</td>
<td></td>
</tr>
<tr>
<td>Smoking Status</td>
<td></td>
<td></td>
<td></td>
<td>0.70</td>
</tr>
<tr>
<td>Current or former</td>
<td>87 (96.7)</td>
<td>39 (97.5)</td>
<td>48 (96.0)</td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>3 (3.3)</td>
<td>1 (2.5)</td>
<td>2 (4.0)</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Histology</td>
<td></td>
<td></td>
<td></td>
<td>0.001</td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>40 (44.0)</td>
<td>25 (61.0)</td>
<td>15 (30.0)</td>
<td></td>
</tr>
<tr>
<td>SCC</td>
<td>38 (41.8)</td>
<td>8 (19.5)</td>
<td>30 (60.0)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>13 (14.3)</td>
<td>8 (19.5)</td>
<td>5 (10.0)</td>
<td></td>
</tr>
<tr>
<td>Adjuvant chemotherapy</td>
<td></td>
<td></td>
<td></td>
<td>0.04</td>
</tr>
<tr>
<td>No</td>
<td>41 (45.1)</td>
<td>13 (31.7)</td>
<td>28 (56.0)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>50 (54.9)</td>
<td>28 (68.3)</td>
<td>22 (44.0)</td>
<td></td>
</tr>
<tr>
<td>Adjuvant radiotherapy</td>
<td></td>
<td></td>
<td></td>
<td>0.98</td>
</tr>
<tr>
<td>No</td>
<td>80 (87.9)</td>
<td>36 (87.8)</td>
<td>44 (88.0)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>11 (12.1)</td>
<td>5 (12.2)</td>
<td>6 (12.0)</td>
<td></td>
</tr>
</tbody>
</table>

TNM denotes tumor-node-metastasis.

* for categorical data, Chi-square test or fisher exact when required, and Wilcoxon signed-rank test for continuous data

** Age at surgery
A B

C D
Clinical Cancer Research

Molecular characteristics of ERCC1 negative vs. ERCC1 positive tumors in resected NSCLC

Luc Friboulet, Daniel Barrios-Gonzales, Frederic Commo, et al.

Clin Cancer Res  Published OnlineFirst July 12, 2011.

Updated version  Access the most recent version of this article at: doi:10.1158/1078-0432.CCR-11-0790

Supplementary Material  Access the most recent supplemental material at: http://clincancerres.aacrjournals.org/content/suppl/2011/07/12/1078-0432.CCR-11-0790.DC1

Author Manuscript  Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions  To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.