DNA Repair Expression Profiling to Identify High-Risk Cytogenetically Normal Acute Myeloid Leukemia and Define New Therapeutic Targets
Ludovic Gabellier, Caroline Bret, Guillaume Bossis, Guillaume Cartron, Jérôme Moreaux

To cite this version:
Ludovic Gabellier, Caroline Bret, Guillaume Bossis, Guillaume Cartron, Jérôme Moreaux. DNA Repair Expression Profiling to Identify High-Risk Cytogenetically Normal Acute Myeloid Leukemia and Define New Therapeutic Targets. Cancers, MDPI, 2020, 12 (10), pp.2874. 10.3390/cancers12102874 . hal-03014671

HAL Id: hal-03014671
https://hal.umontpellier.fr/hal-03014671
Submitted on 1 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
DNA Repair Expression Profiling to Identify High-Risk Cytogenetically Normal Acute Myeloid Leukemia and Define New Therapeutic Targets

Ludovic Gabellier 1,2,3, Caroline Bret 2,4,*, Guillaume Bossis 5,5, Guillaume Cartron 1,2,3 and Jérôme Moreaux 2,4,6,7,*

1 Département d’Hématologie Clinique, CHU Montpellier, University of Montpellier, Montpellier, France
2 UFR de Médecine, University of Montpellier, Montpellier, France
3 Institut de Génétique Moléculaire de Montpellier (IGMM), University of Montpellier, CNRS, Montpellier, France
4 CHU Montpellier, Department of Biological Hematology, Montpellier, France
5 Equipe Labellisée Ligue Contre le Cancer, Paris, France
6 Institute of Human Genetics, IGH, CNRS, Univ Montpellier, France
7 Institut Universitaire de France (IUF)

* Correspondence: jerome.moreaux@igh.cnrs.fr

Received: date; Accepted: date; Published: date

Abstract: Cytogenetically normal acute myeloid leukemias (CN-AML) represent about 50% of total adult AML. Despite the well-known prognosis role of gene mutations such as NPM1 mutations or FLT3 internal tandem duplication (FLT3-ITD), clinical outcomes remain heterogeneous in this subset of AML. Given the role of genomic instability in leukemogenesis, expression analysis of DNA repair genes might be relevant to sharpen prognosis evaluation in CN-AML. Publicly available gene expression profile dataset from two independent cohorts of patients with CN-AML were analyzed (GSE12417). We investigated the prognostic value of 175 genes involved in DNA repair. Among these genes, 23 were associated with a prognostic value. The prognostic information provided by these genes was summed in a DNA repair score to consider connection of DNA repair pathways. DNA repair score allowed to define a group of patients (n=87; 53,7%) with poor median overall survival (OS) of 233 days (95% CI: 184-260). These results were confirmed in the validation cohort (median OS: 120 days; 95% CI: 36-303). In multivariate Cox analysis, the DNA repair score, NPM1 and FLT3-ITD mutational status remained independent prognosis factors in CN-AML. Combining these parameters allowed the identification of three risk groups with different clinical outcomes in both training and validation cohorts. Combined with NPM1 and FLT3 mutational status, our GE-based DNA repair score might be used as a biomarker to predict outcomes for patients with CN-AML. DNA repair score has the potential to identify CN-AML patients whose tumor cells are dependent on specific DNA repair pathways to design new therapeutic avenues.

Keywords: acute myeloid leukemia; normal karyotype; DNA repair; risk score; precision medicine

1. Introduction

Acute myeloid leukemia (AML) is the most frequent type of adult leukemia. When analyzed with conventional cytogenetics, about 40-50% of AML exhibit no chromosomal abnormalities, and are defined as “cytogenetically normal AML” (CN-AML)[1]. Recurrent mutated genes in CN-AML were identified, such as NPM1, signal transduction genes (FLT3) or myeloid transcription factor genes (CEBPA, RUNX1)[2]. Based on presence, absence and allelic ratio of these mutations, CN-AML may be classified in favorable, intermediate or adverse prognosis, illustrating the high heterogeneity of clinical outcomes in this AML subset[3]. Yet, a wide diversity of gene mutations occurring in...
CN-AML were revealed by deep sequencing techniques, such as mutations of DNA modification, cohesin or tumor-suppressor genes, suggesting the wide heterogeneity of molecular mechanisms involved in leukemogenesis[4-6].

Even if the study of mutational landscape by new DNA sequencing technologies demonstrated a low mutation frequency in AML compared to others cancers[7], genomic instability remains a well-described leukemogenesis mechanism, illustrated by the high frequency of AML with non-random cytogenetics abnormalities or with complex karyotype[8, 9]. Therefore, the role of DNA damage response (DDR) in the AML field has been widely studied. Polymorphic variants of genes involved in several DNA repair pathways had been associated with the onset of AML, such as XPD-Lys751Gln, involved in the nucleotide excision repair mechanism[10]. Recurrent AML fusion transcripts such as RUNXI-RUNXIT1 or PML-RARA has also been demonstrated to downregulate the expression of genes implied in DDR[11-14]. Moreover, children or young adults AML are often associated with hereditary diseases due to DNA repair gene mutations, such as Fanconi disease[15], Bloom syndrome or Werner syndrome[16]. Finally, dysregulation in DDR also contribute to increased resistance to conventional chemotherapy by several mechanisms, such as paradoxical increased expression of DDR or cell cycle check-point genes[17-19].

In the current study, we investigate the prognostic value of genes related to the major DNA repair pathways. The data reveals specific patterns of gene expression in CN-AML that have prognostic value. Therefore, the expression analysis of DNA repair genes might be relevant in the context of CN-AML to sharpen prognosis evaluation of this heterogeneous AML subset.

2. Results

2.1. Linking Expression of DNA Repair Genes and AML Patient Overall Survival

Considering the important role of DNA repair in drug resistance and adaptation to replication stress in cancer cells, we first aimed to identify the DNA repair genes associated with overall survival in CN-AML. A list set of 175 genes involved in six major DNA repair pathways (base excision repair (BER), NER, mismatch repair (MMR), homologous recombination repair (HRR), non-homologous end joining (NHEJ) and FANC pathways) was defined using the REPAIRtoire database (http://repairtoire. genesilico.pl) and review of the literature (Supplementary Table S1). Using the MaxStat R function, we identified 23 out of the 175 genes which level of expression had a prognostic value in the two independent cohorts. Nineteen genes were associated with poor prognosis and 4 genes with good prognosis (Table 1). No statistically significant prognostic value was found for any gene involved in NHEJ pathway.

<table>
<thead>
<tr>
<th>DNA repair pathway</th>
<th>Probe set</th>
<th>Gene symbol</th>
<th>Benjamini Hochberg corrected p-value</th>
<th>Hazard ratio</th>
<th>Prognosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base Excision</td>
<td>210027_s_at</td>
<td>APEX1</td>
<td>0.02</td>
<td>1.6</td>
<td>Bad</td>
</tr>
<tr>
<td>Repair pathway</td>
<td>209731_at</td>
<td>NTHL1</td>
<td>0.0016</td>
<td>1.9</td>
<td>Bad</td>
</tr>
<tr>
<td>(BER)</td>
<td>202330_s_at</td>
<td>UNG</td>
<td>0.0095</td>
<td>2</td>
<td>Bad</td>
</tr>
<tr>
<td></td>
<td>203655_at</td>
<td>XRCC1</td>
<td>0.022</td>
<td>1.6</td>
<td>Bad</td>
</tr>
<tr>
<td>Fanconi</td>
<td>209902_at</td>
<td>ATR</td>
<td>0.0048</td>
<td>1.8</td>
<td>Bad</td>
</tr>
<tr>
<td>pathway</td>
<td>214727_at</td>
<td>BRCA2</td>
<td>0.0049</td>
<td>0.58</td>
<td>Good</td>
</tr>
<tr>
<td>(FANC)</td>
<td>203719_at</td>
<td>ERCC1</td>
<td>0.0037</td>
<td>1.9</td>
<td>Bad</td>
</tr>
<tr>
<td></td>
<td>203678_at</td>
<td>FAN1</td>
<td>0.0028</td>
<td>1.8</td>
<td>Bad</td>
</tr>
<tr>
<td></td>
<td>221206_at</td>
<td>PMS2 ///</td>
<td>0.024</td>
<td>1.8</td>
<td>Bad</td>
</tr>
<tr>
<td></td>
<td>219317_at</td>
<td>PMS2CL</td>
<td>0.0016</td>
<td>1.9</td>
<td>Bad</td>
</tr>
<tr>
<td>Homologous</td>
<td>214727_at</td>
<td>BRCA2</td>
<td>0.0049</td>
<td>0.58</td>
<td>Good</td>
</tr>
<tr>
<td>Recombination</td>
<td>205395_s_at</td>
<td>MRE11A</td>
<td>0.015</td>
<td>1.8</td>
<td>Bad</td>
</tr>
<tr>
<td>Repair</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
To further corroborate gene expression data on a functional level, we studied CRISPR or RNAi screening publicly available data (Dependency Map data, Broad Institute, www.depmap.org) [20, 21]. Interestingly, among the 19 genes associated with a poor outcome, APEX1 (BER), RTEL1 (HRR) and COPS6 (NER) were identified as significant essential AML genes (p = 7.9e-05, 3.4e-04 and 2.8e-04 respectively) (Figure 1).

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Gene</th>
<th>DEMETER2 Effect</th>
<th>p-value</th>
<th>Cell Line</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mismatch Repair</strong> (MMR)</td>
<td>MSH3</td>
<td>0.000043</td>
<td>2.8</td>
<td>Bad</td>
</tr>
<tr>
<td></td>
<td>MSH2</td>
<td>0.024</td>
<td>1.8</td>
<td>Bad</td>
</tr>
<tr>
<td></td>
<td>RFC2</td>
<td>0.023</td>
<td>1.6</td>
<td>Bad</td>
</tr>
<tr>
<td><strong>Nucleotide Excision Repair</strong> (NER)</td>
<td>COP56</td>
<td>0.011</td>
<td>1.7</td>
<td>Bad</td>
</tr>
<tr>
<td></td>
<td>EP300</td>
<td>0.019</td>
<td>0.59</td>
<td>Good</td>
</tr>
<tr>
<td></td>
<td>ERCC1</td>
<td>0.0037</td>
<td>1.9</td>
<td>Bad</td>
</tr>
<tr>
<td></td>
<td>ERCC8</td>
<td>0.04</td>
<td>1.5</td>
<td>Bad</td>
</tr>
<tr>
<td></td>
<td>GTF2H2</td>
<td>0.033</td>
<td>1.5</td>
<td>Bad</td>
</tr>
<tr>
<td></td>
<td>RAD23A</td>
<td>0.0067</td>
<td>0.53</td>
<td>Good</td>
</tr>
<tr>
<td></td>
<td>XPA</td>
<td>0.0035</td>
<td>1.8</td>
<td>Bad</td>
</tr>
<tr>
<td></td>
<td>XRCC1</td>
<td>0.022</td>
<td>1.6</td>
<td>Bad</td>
</tr>
</tbody>
</table>

**APEX1**

**RTEL1**

**COPS6**

Figure 1. Silencing of APEX1, RTEL1 and COPS6 impairs AML cell growth. Using CRISPR or RNAi screening publicly available data (Dependency Map data, Broad Institute, www.depmap.org).
dependency scores of APEXI, RTEL1 and COPS6 underline their specific importance for AML cell survival compared to all cell lines tested.

2.2. GEP-Based DNA Repair Score for Predicting CN-AML Patients’ Survival

Then, we searched to combine the prognostic information of these genes in a GE-based DNA repair risk score. The 23 DNA repair genes associated with a prognostic value included 4 coding genes for BER pathway, 6 genes for FANC pathway, 6 genes for HRR pathway, 3 genes for MMR pathway and 8 genes for NER pathway (Table 1). Four out of these 23 probesets (BRCA2, ERCC1, PMS2//PMS2CL and XRCC1) were involved in two different pathways. A specific GE-based risk score was established for BER, FANC, HRR, MMR and NER DNA repair pathways. GE-based DNA repair scores were defined by the sum of the beta coefficients of the Cox model for each prognostic gene, weighted by +1 or -1 according to the patient signal above or below / equal the probe set MaxStat value as previously described[22, 23]. Using Maxstat R function, high BER, FANC, HRR, MMR and NER score values were significantly associated with poor prognosis in the training cohort (Supplementary Figure S1).

In Cox multivariate analysis, only HRR and NER scores remained associated with overall survival in the training cohort (Table 2). Therefore, a global DNA repair score was established, incorporating the prognostic value of HRR and NER scores. To this aim, CN-AML patients were split in three subgroups: group I included patients with low NER and HRR risk score values (n=20), group III included patients with high NER and HRR risk scores (n=87) and group II included patients with NER or HRR high-risk score value (n=55).

Table 2. Cox analysis of overall survival in CN-AML training cohort (n=162) according to DNA repair pathway scores. Hazard ratio (HR) and p-values are shown for each DNA repair pathway score in univariate and multivariate Cox analysis. NS: not significant.

<table>
<thead>
<tr>
<th>DNA repair pathway score</th>
<th>Univariate Cox analysis</th>
<th>Multivariate Cox analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR</td>
<td>p-value</td>
</tr>
<tr>
<td>BER score</td>
<td>1.97</td>
<td>1.44e-03</td>
</tr>
<tr>
<td>FANC score</td>
<td>2.32</td>
<td>2.98e-05</td>
</tr>
<tr>
<td>HRR score</td>
<td>3.23</td>
<td>2.16e-07</td>
</tr>
<tr>
<td>MMR score</td>
<td>2.80</td>
<td>1.59e-04</td>
</tr>
<tr>
<td>NER score</td>
<td>3.83</td>
<td>2.90e-04</td>
</tr>
</tbody>
</table>

After a median follow-up of 1176 days (95% CI: 916-969), median overall survival (OS) was 293 days (95% CI: 252-461) for the whole training cohort (Supplementary Figure S2a). One-year OS was 45.2% (95% CI: 38.0-53.8). According to risk groups determined by the DNA repair score, median OS was not reached (95% CI: NR-NR), 693 days (95% CI: 414-NR) and 233 days (95% CI: 184-260) respectively for patients in groups I, II and III (Figure 2a). Median OS were statistically different between each risk group (log-rank test; p = 0.016 between group I and II; p < 0.001 between group II and III).

We searched to validate these results in an independent cohort of 78 patients. HRR and NER scores computed with training cohort parameters were also prognostic in this validation cohort (Supplementary Table S2). The global DNA repair score was also computed. In the validation set, risk groups included 14, 42 and 22 patients respectively in groups I, II and III. After a median follow-up of 1183 days (95% CI: 1092-1383), median overall survival (OS) was 538 days (95% CI: 388-1278) for the whole validation cohort (Supplementary Figure S2b). One-year OS was 61.1% (95% CI: 51.1-73.0). According to risk groups determined by the DNA repair score, median OS was not reached (95% CI: 538-NR), 787 days (95% CI: 473-NR) and 120 days (95% CI: 36-303) respectively for patients in groups I, II and III (Figure 2b). Even if survival analysis failed to demonstrate a statistical difference between groups I and II (log-rank test; p = 0.287), OS was still statistically different between risk groups II and III (log-rank test; p < 0.001). Altogether, these data underlined the identification of high-risk CN-AML patients characterized by DNA repair dysregulation and that could benefit from DNA repair targeted treatment.
Figure 2. Kaplan-Meier survival curves according to risk stratification determined by DNA repair score. (a) Kaplan-Meier survival curve in the training cohort (n=162). Median OS was not reached (95% CI: NR-NR), 693 days (95% CI: 414-NR) and 233 days (95% CI: 184-260) respectively for patients in groups I (low DNA repair score), II (medium DNA repair score) and III (high DNA repair score). One-year OS was 90.0% (95% CI: 77.7-100) in group I, 62.8% (95% CI: 51.1-77.2) in group II, and 23.4% (95% CI: 15.8-34.7) in group III. (b) Kaplan-Meier survival curve in the validation cohort (n=78). Median OS was not reached (95% CI: 538-NR), 787 days (95% CI: 473-NR) and 120 days (95% CI: 36-303) respectively for patients in groups I (low DNA repair score), II (medium DNA repair score) and III (high DNA repair score). One-year OS was 85.7% (95% CI: 69.2-100) in group I, 73.3% (95% CI: 60.9-88.2) in group II, and 22.7% (95% CI: 10.5-49.1) in group III. P-values are determined with log-rank test. NR: not reached.
2.3. DNA Repair Score and NPM1 / FLT3 Mutational Status Combination as Prognosis Factors in CN-AML

Because NPM1 mutations and FLT3-ITD (internal tandem duplications) are well-described prognosis factors in CN-AML, we conducted another Cox analysis to determine whether our DNA repair score provides additional prognostic information. Prognostic classification according to NPM1 and FLT3 mutational status was established in both cohort according to actual recommendations[3]: patients with only NPM1 mutation were classified as “better outcome”, patients with only FLT3-ITD were classified as “adverse prognosis” and patients with both or none of these mutations were classified as “intermediate prognosis”. Kaplan-Meier survival curves according to NPM1 and FLT3 mutational status are presented in Supplementary Figure S3 for both cohorts.

Using multivariate Cox analysis, our DNA repair score and NPM1/FLT3 mutation classification remained independently associated with survival (Table 3 & Supplementary Table S3). Therefore, we investigated the interest of combining DNA repair score and NPM1 / FLT3 mutational status to predict CN-AML outcome. Patients were classified according to prognosis value of DNA repair score (0 point for group I; 1 for group II; 2 for group III), and NPM1 / FLT3 mutational status (0 point if NPM1 mutated without FLT3-ITD; 2 points if FLT3-ITD without NPM1 mutation; 1 point in other situations). The sum of the prognostic information was computed for all patients, allowing to separate patients in three new prognostic groups: group A including patients with 0 or 1 point, group B for patients with 2 points and group C for patients with 3 or 4 points. (Table 4).

### Table 3. Cox analysis of overall survival in CN-AML training cohort (n=162) according to DNA repair score, and NPM1 & FLT3 mutational status. Hazard ratio (HR) and p-values are shown for each parameter in univariate and multivariate Cox analysis. ITD: internal tandem duplication.

<table>
<thead>
<tr>
<th>Scores</th>
<th>Univariate Cox analysis</th>
<th>Multivariate Cox analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR</td>
<td>p-value</td>
</tr>
<tr>
<td>DNA repair score</td>
<td>2.76</td>
<td>1.49e-08</td>
</tr>
<tr>
<td>NPM1 mutation / FLT3-ITD classification</td>
<td>1.81</td>
<td>1.18e-04</td>
</tr>
</tbody>
</table>

### Table 4. DNA repair score and NPM1 / FLT3 mutational status combination in order to establish a global prognosis score in CN-AML. Patients were classified according to DNA repair score risk group (I, II or III) and NPM1 / FLT3 mutational status. Patients with NPM1 mutation and FLT3-ITD are respectively designated by NPM1+ and FLT3-ITD+. Patients without NPM1 mutation or FLT3-ITD are respectively designated by NPM1- and FLT3-ITD-. Points were attributed as described in the table. Patients with 0 or 1 point were grouped in group A (green), patients with 2 points were grouped in group B (yellow), and patients with 3 or 4 points were grouped in group C (red). ITD: internal tandem duplication.

<table>
<thead>
<tr>
<th>Classification according to DNA repair score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group I</td>
</tr>
<tr>
<td>0 point</td>
</tr>
</tbody>
</table>

NPM1 and FLT3 mutational status

- NPM1+ and FLT3-ITD- 0 point
- NPM1+ and FLT3-ITD+ or NPM1- and FLT3-ITD- 1 point
- NPM1- and FLT3-ITD+ 2 points

In the training cohort, median OS was not reached (95% CI: NR-NR), 326 days (95% CI: 127-NR) and 236 days (95% CI: 190-263) respectively for patients in groups A, B and C. One-year OS was 90.3% (95% CI: 80.5-100) in group A, 49.3% (95% CI: 37.1-65.7) in group B, and 24.2% (95% CI: 16.2-36.2) in group C. These results were confirmed in the validation cohort where median OS was not reached (95% CI: 1278-NR), 516 days (95% CI: 308-NR) and 253 days (95% CI: 52-403) for patients...
respectively in groups A, B and C. One-year OS was 92.6% (95% CI: 83.2-100) in group A, 54.9% (95% CI: 39.8-75.7) in group B, and 26.5% (95% CI: 12.4-55.8) in group C. OS was statistically different between groups A, B and C in both training and validation cohorts (Figure 3). Altogether, these data underlined the interest of GEP-based DNA repair deregulations, alone or in combination with NPM1 and FLT3 mutational status to identify high-risk CN-AML patients.

Figure 3. Kaplan-Meier survival curves according to risk groups determined by combined score incorporating DNA repair score and NPM1/FLT3 mutational status. (a) Kaplan-Meier survival curve in the training cohort (n=162). Median OS was not reached (95% CI: NR-NR), 326 days (95% CI: 127-NR) and 236 days (95% CI: 190-263) respectively for patients in groups A, B and C. One-year OS was 90.3% (95% CI: 80.5-100) in group A, 49.3% (95% CI: 37.1-65.7) in group B, and 24.2% (95% CI: 16.2-36.2) in group C.

(b) Kaplan-Meier survival curve in the validation cohort (n=78). Median OS was not reached (95% CI:}
3. Discussion

Despite improvement in prognosis classification, mostly based on the identification of gene mutations such as NPM1, FLT3 or CEBPA, outcomes in CN-AML remain heterogeneous, underlying the wide diversity of this AML subset. In this study, we developed a GE-based score using data from genes involved in DNA damage response. Our model succeeded to predict poor outcomes in two independent cohorts of adult patients with CN-AML treated with intensive chemotherapy. Combining DNA repair score with NPM1 and FLT3-ITD mutational status allows to distinguish three prognostic groups including a low-risk group with a not reached median OS after a median follow-up of more than 3 years in both cohorts, a high-risk group with a median OS of about 8 months in both cohorts, and an intermediate risk-group. This model may therefore be used for risk stratification in CN-AML.

Among the GEP-based defined DNA-repair scores built in our study, HRR and NER scores remained independent prognostic factors in CN-AML. HRR pathway is a process involved in DNA double-strand break (DSB) repair, in which complementary sister chromatid is used as a template for an error-free repair of DNA sequence[24, 25]. Among the prognostic factors composing the DNA repair score, MRE11A is a nucleosome involved in the MRN complex (for MRE11 - RAD50 - NBS1) which acts as a sensor for DSB damage[26, 27]. RAD52, BRCA2, XRCC2 are proteins directly involved in the DNA repair process[25, 28], and RTEL1 and SRCAp are regulators of HRR[29, 30]. NER pathway is involved in recognition and repair of lesions that disrupt DNA double helix, such as adducts or inter-strand crosslinks (ICL)[31, 32]. RAD23A and COPS6 are involved in DNA damage recognition. The recruitment of the DNA incision complex, in which ERCC1, ERCC8 and GTF2H2 are involved, is mediated by XPA[31, 33]. XRCC1 and EP300 are respectively involved in DNA final ligation process and NER regulation[34, 35]. Several polymorphisms in genes involved in HRR and NER have been correlated with AML onset and outcome. RAD51 is a key protein in HRR pathway. Its polymorphic variant RAD51-G135C has been suggested to be correlated with the onset of therapy-related AML by several case-control studies, even if two meta-analysis seem to dismiss the role of this polymorphism in de novo AML onset[36-39]. XPD is involved in NER pathway, and its polymorphism XPD-Lys751Gln has been shown to be a risk factor for AML onset[10, 36, 37]. One study also suggested that this polymorphism worsens the AML prognosis[40]. These data highlight the role of DNA repair pathways in leukemogenesis, and suggest their resistance in chemotherapy resistance.

Interestingly, when compared using multivariate analysis, the DNA repair score and NPM1/FLT3 mutational status remained statistically associated with outcome in CN-AML. FLT3 and NPM1 have also been shown to play a role in DNA damage response in AML. FLT3-ITD mutations, occurring in about 20-25% of CN-AML, leads to a constitutive activation of FLT3, and therefore confers a growth advantage to leukemic cells. Several studies showed that the level of reactive oxygen species (ROS) was increased in FLT3-ITD mutated AML cells, and correlated with high levels of DSB and lower efficiency of NHEJ repair pathway[41]. Moreover, the use of tyrosine-kinase inhibitors may reduce both ROS and DSB levels, and increase DNA repair efficiency, overcoming the chemo-resistance of these cells[41, 42]. Other mechanisms have been suggested to explain the role of FLT3-ITD in DNA damages and acquired drug resistance of AML cells, such as telomere-related genome instability[43], or paradoxical up-regulation of RAD51[44]. NPM1 is the most commonly mutated gene in CN-AML, with more than 50 described mutations. The prognostic significance of these mutations and co-mutations in other genes has been widely studied[45]. The role of NPM1 in DNA damage response and maintenance of genome stability is less clear. NPM1 is involved in regulation of centrosome duplication during cell cycle[46], or is recruited in its phosphorylated form (NPM1-pT199) on DSB foci, even if its role in DSB repair remains discussed[47]. NPM1 is also involved in regulation of key DNA repair factors, such as APEX1 or p53[48, 49]. Therefore, NPM1
mutations in AML result in APEX1 abnormal cytoplasmic accumulation, and impaired BER activity[50], potentially explaining a chemotherapy improved response in NPM1-mutated AML.

Intensive chemotherapy for CN-AML patients usually includes cytarabine and anthracyclines (daunorubicine or idarubicine)[51]. Cytarabine, a nucleoside analog, incorporates into DNA and interferes with DNA synthesis during the phase S of the cell cycle, leading to genomic instability[52]. Anthracyclines are DNA topoisomerase II inhibitors that induce DNA damages such as DSB, adducts and ICL[52]. Therefore, overexpression of HRR or NER pathway genes could be associated with chemotherapy resistance, but a better understanding of the functional role of DNA repair pathways in the pathogenesis and drug resistance of CN-AML is needed[53]. Gene silencing approaches by shRNA or CRISPR-Cas9 strategies could be of particular interest. Of particular interest, CRISPR-Cas9 or RNAi screening revealed that APEX1 (BER), RTEL1 (HRR) and COPS6 (NER) are essential AML genes. Among these genes, COPS6 overexpression is associated with poor outcome in many solid tumors. Interestingly, COPS6 depletion showed in vivo efficacy against glioblastoma[54], cervical cancer[55] or papillary thyroid carcinoma[56], through regulation of several signaling pathways. However, the biological function of COPS6 in leukemogenesis and AML drug-resistance, remains largely unknown.

Therefore, inhibiting DNA repair might be a promising strategy to improve the efficacy of genotoxic drugs and overcome drug resistance, according to the principle of "synthetic lethality"[57, 58]. APEX1 inhibitor has demonstrated a promising toxicity on primary AML cells in vitro, alone or in association with hypomethylating agent decitabine or PARP (poly(ADP-Ribose) polymerase) inhibitor talazoparib. Even if APEX1 expression levels did not significantly differ between responding and non-responding AML cells, APEX1 inhibitor appeared promising in normal karyotype AML (83% of the APEX1 inhibitor “responders”) [59]. Our data support the potential therapeutic interest of DNA damage signaling and DNA repair inhibitors in CN-AML.

4. Materials and Methods

4.1. Patients and Gene Expression Data

Gene expression microarray data from two independent cohorts of adult patients diagnosed with CN-AML were used. The first cohort (training set) included 162 patients and the second one (validation set) 78 patients. At least 20 metaphases were analyzed for each patient to confirm the normal karyotype. At the beginning of treatment, median age was 58 years in the training cohort and 62 years in the validation cohort. Pretreatment clinical characteristics of patients have been described previously[60]. NPM1 and FLT3 mutational status were kindly provided for each patient by Metzeler et al[60]. All patients were treated with intensive chemotherapy.

Affymetrix gene expression data are publicly available via the online Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE12417. They were performed using Affymetrix HG-U133 A&B microarrays for first cohort and Affymetrix HG-U133 P 2.0 microarrays for the second one. Normalization of microarray data was performed using the variance stabilizing normalization algorithm, and probe set signals calculated by the median polish method[60, 61].

Quality control consisted of visual inspection of the array image for artifacts, assessment of RNA degradation plots, and inspection of rank-vs-residual plots after normalization and probe set summarization.

4.2. Selection of Prognostic Genes

DNA repair gene list was defined using the REPAIRtoire database (http://repairtoire.genesilico.pl) and review of the literature (Supplementary Table S1)[62]. To establish gene expression (GE)-based risk scores, we selected probe sets whose expression values were significantly associated with overall survival, using MaxStat R function and Benjamini Hochberg multiple testing correction (adjusted p-value < 0.05)[22].

4.3. Building DNA Repair Gene Expression-Based Risk Score
For each pathway, a GE-based risk score was created as the sum of the beta coefficients weighted by +1 or -1 according to the patient signal above or below / equal the probe set MaxStat value as previously reported[22, 23]. Patients from the training cohort were ranked according to increased prognostic score and for a given score value X, the difference in survival of patients with a prognostic score ≤X or >X was computed using MaxStat analysis.

Cox proportional hazards model was performed to determine statistically significant pathway scores in multivariate analysis. A global DNA repair score was calculated based on the pathway scores which remained statistically significant in this analysis. Survival analyses were assessed using Kaplan-Meier method, and survival curves were compared using log-rank test.

4.4. Validation of the DNA Repair Score on Validation Cohort

Pathway and DNA repair scores were individually calculated in the validation cohort, using the cutoff values determined for the training cohort. Survival analyses were assessed using Kaplan-Meier method, and survival curves were compared using log-rank test.

4.5. Statistical Analyses

All statistical tests were two-tails and Alpha-risk was fixed at 5%. Analyses were performed using R.3.6.0. and SPSS Statistics version 23.0.0.0 for Mac.

5. Conclusions

The DNA repair score may be useful to identify high-risk CN-AML patients and define the best DNA repair inhibitor to use in combination with conventional treatment to improve patients' outcome. The DNA repair score could also be valuable for adapting targeted treatment according to the drug resistance mechanisms selected during clonal evolution of relapsing AML. These advances may improve the survival of CN-AML patients, and limit the side effects of treatment, improving compliance with dosing regimens and overall quality of life.

Author Contributions: LG performed research, data analyses and participated in the writing of the paper. GB and GC participated in the research and in the writing of the paper. JM and CB supervised the research and the writing of the paper.

Funding: This work was supported by grants from INCa (Institut National du Cancer) PLBIO18-362 PIT-MM and PLBIO19 FATidique, ANR (TIE-Skip; 2017-CE15-0024-01), ANR-18-CE15-0010-01 PLASMADIFF-3D, SIRIC Montpellier Cancer (INCa_Inserm_DGOS_12553), Labex EpiGenMed and Institut Universitaire de France.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A - Supplementary Tables & Figures

Supplementary Table S1. Genes coding for proteins involved in DNA repair. Gene symbols are provided with corresponding probe sets for each DNA repair pathway.
### Fanconi (FANC) pathway

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>213454_at</td>
<td>APITD1</td>
</tr>
<tr>
<td>208442_s_at</td>
<td>ATM</td>
</tr>
<tr>
<td>209902_s_at</td>
<td>ATR</td>
</tr>
<tr>
<td>1552937_s_at</td>
<td>ATRIP</td>
</tr>
<tr>
<td>205733_s_at</td>
<td>BLM</td>
</tr>
<tr>
<td>204531_s_at</td>
<td>BRCA1</td>
</tr>
<tr>
<td>214727_at</td>
<td>BRCA2</td>
</tr>
<tr>
<td>221800_s_at</td>
<td>C17orf70</td>
</tr>
<tr>
<td>214816_x_at</td>
<td>C19orf40</td>
</tr>
<tr>
<td>208442_s_at</td>
<td>ATM</td>
</tr>
<tr>
<td>205394_at</td>
<td>ATRIP</td>
</tr>
<tr>
<td>203229_s_at</td>
<td>CLK2</td>
</tr>
<tr>
<td>234464_s_at</td>
<td>EME1</td>
</tr>
<tr>
<td>203719_at</td>
<td>ERCC1</td>
</tr>
<tr>
<td>228131_at</td>
<td>EME1</td>
</tr>
<tr>
<td>204531_s_at</td>
<td>BRCA1</td>
</tr>
<tr>
<td>214727_at</td>
<td>BRCA2</td>
</tr>
<tr>
<td>221800_s_at</td>
<td>C17orf70</td>
</tr>
<tr>
<td>214816_x_at</td>
<td>C19orf40</td>
</tr>
<tr>
<td>208442_s_at</td>
<td>ATM</td>
</tr>
<tr>
<td>205394_s_at</td>
<td>ATRIP</td>
</tr>
<tr>
<td>203229_s_at</td>
<td>CLK2</td>
</tr>
<tr>
<td>234464_s_at</td>
<td>EME1</td>
</tr>
<tr>
<td>203719_at</td>
<td>ERCC1</td>
</tr>
<tr>
<td>228131_at</td>
<td>EME1</td>
</tr>
<tr>
<td>204531_s_at</td>
<td>BRCA1</td>
</tr>
<tr>
<td>214727_at</td>
<td>BRCA2</td>
</tr>
<tr>
<td>221800_s_at</td>
<td>C17orf70</td>
</tr>
<tr>
<td>214816_x_at</td>
<td>C19orf40</td>
</tr>
</tbody>
</table>

### Homologous Recombination Repair (HRR) pathway

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>208442_s_at</td>
<td>ATM</td>
</tr>
<tr>
<td>205345_at</td>
<td>BARD1</td>
</tr>
<tr>
<td>205733_at</td>
<td>BLM</td>
</tr>
<tr>
<td>204531_at</td>
<td>BRCA1</td>
</tr>
<tr>
<td>214727_at</td>
<td>BRCA2</td>
</tr>
<tr>
<td>221800_s_at</td>
<td>C17orf70</td>
</tr>
<tr>
<td>214816_x_at</td>
<td>C19orf40</td>
</tr>
<tr>
<td>208442_s_at</td>
<td>ATM</td>
</tr>
<tr>
<td>205394_s_at</td>
<td>ATRIP</td>
</tr>
<tr>
<td>203229_s_at</td>
<td>CLK2</td>
</tr>
<tr>
<td>234464_s_at</td>
<td>EME1</td>
</tr>
<tr>
<td>203719_at</td>
<td>ERCC1</td>
</tr>
<tr>
<td>228131_at</td>
<td>EME1</td>
</tr>
<tr>
<td>204531_s_at</td>
<td>BRCA1</td>
</tr>
<tr>
<td>214727_at</td>
<td>BRCA2</td>
</tr>
<tr>
<td>221800_s_at</td>
<td>C17orf70</td>
</tr>
<tr>
<td>214816_x_at</td>
<td>C19orf40</td>
</tr>
</tbody>
</table>

### Mismatch Repair (MMR) pathway

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>204603_at</td>
<td>EXO1</td>
</tr>
<tr>
<td>202726_at</td>
<td>LIG1</td>
</tr>
<tr>
<td>202520_s_at</td>
<td>MLH1</td>
</tr>
<tr>
<td>204838_s_at</td>
<td>MLH3</td>
</tr>
<tr>
<td>209421_at</td>
<td>MSH2</td>
</tr>
<tr>
<td>205887_s_at</td>
<td>MSH3</td>
</tr>
<tr>
<td>202911_at</td>
<td>MSH6</td>
</tr>
<tr>
<td>201202_at</td>
<td>PCNA</td>
</tr>
</tbody>
</table>

### Nucleotide Excision Repair (NER) pathway

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>204093_at</td>
<td>CCNH</td>
</tr>
<tr>
<td>211297_s_at</td>
<td>CDK7</td>
</tr>
<tr>
<td>209194_at</td>
<td>CETN2</td>
</tr>
<tr>
<td>2042467_s_at</td>
<td>COPS2</td>
</tr>
<tr>
<td>202078_at</td>
<td>COPS3</td>
</tr>
<tr>
<td>218042_at</td>
<td>COPS4</td>
</tr>
<tr>
<td>201652_s_at</td>
<td>COPS5</td>
</tr>
<tr>
<td>201405_s_at</td>
<td>COPS6</td>
</tr>
</tbody>
</table>
Supplementary Table S2. Cox analysis of overall survival in CN-AML validation cohort (n=78) according to DNA repair pathway scores. Hazard ratio (HR) and p-values are shown for each HRR and NER repair pathway score (computed with training cohort parameters) in univariate Cox analysis.

<table>
<thead>
<tr>
<th>Scores</th>
<th>Univariate Cox analysis</th>
<th>Multivariate Cox analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA repair pathway score</td>
<td>HR 3.73 p=1.32e-05</td>
<td>HR 3.04 p=1.01e-05</td>
</tr>
<tr>
<td>HRR score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NER score</td>
<td>2.83 p=0.028</td>
<td></td>
</tr>
</tbody>
</table>

Supplementary Table S3. Cox analysis of overall survival in CN-AML validation cohort (n=78) according to DNA repair score, and NPM1 & FLT3 mutational status. Hazard ratio (HR) and p-values are shown for each parameter in univariate and multivariate Cox analysis. NS: not significant. ITD: internal tandem duplication.

<table>
<thead>
<tr>
<th>Scores</th>
<th>Univariate Cox analysis</th>
<th>Multivariate Cox analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA repair score</td>
<td>3.04 p=1.01e-05</td>
<td>3.07 p=1.4e-05</td>
</tr>
<tr>
<td>NPM1 mutation / FLT3-ITD classification</td>
<td>1.71 p=0.020</td>
<td>1.67 p=0.03</td>
</tr>
</tbody>
</table>

a)
Cancers 2020, 12, x FOR PEER REVIEW

338 339

b) Maimaly selected rank statistics

Standardized log-rank statistic

\[ \text{FANC score cut-point: } -0.57621 \]

\[ P = 1.79 \times 10^{-5} \]

340 341 342
c) Maimaly selected rank statistics

Standardized log-rank statistic

\[ \text{HRR score cut-point: } -0.26745 \]

\[ P = 5.08 \times 10^{-8} \]

343 344 345
d) Maimaly selected rank statistics

Standardized log-rank statistic

\[ \text{MMRAB cut-point: } 0.012234 \]

\[ P = 8.01 \times 10^{-5} \]

346 347 348
e)
Supplementary Figure S1. Prognostic value of DNA repair pathway scores in CN-AML patients of the training cohort. Patients of the training cohort (n=162) were ranked according to increasing BER (a), FANC (b), HRR (c), MMR (d) and NER (e) scores and a maximum difference in OS was obtained using MaxStat R function. Green survival curves represent patients whose score is inferior or equal to the MaxStat determined cut-point. Red survival curves designate patients whose score is strictly superior to the MaxStat determined cut-point.
Supplementary Figure S2. Kaplan-Meier survival curves for training and validation cohorts. (a) Kaplan-Meier survival curve for the whole training cohort (n=162). After a median follow-up of 1176 days (95% CI: 916-NR), median overall survival (OS) was 293 days (95% CI: 252-461) for the whole training cohort. (b) Kaplan-Meier survival curve for the whole validation cohort (n=78). After a median follow-up of 1183 days (95% CI: 1092-1383), median overall survival (OS) was 538 days (95% CI: 388-1278) for the whole validation cohort.
Supplementary Figure S3. Kaplan-Meier survival curves according to NPM1/FLT3 mutational status. (a) Kaplan-Meier survival curve for the training cohort (n=162). Median OS was not reached (95% CI: 999-NR) for patients with NPM1+/FLT3−ITD− mutational status, 271 days (95% CI: 240-416) for patients with NPM1+/FLT3−ITD+ or NPM1−/FLT3− mutational status (“Others”) and 214 days (95% CI: 123-657) for patients with NPM1−/FLT3−ITD+ mutational status. (b) Kaplan-Meier survival curve for the validation cohort (n=78). Median OS was not reached (95% CI: 624-NR) for patients with NPM1+/FLT3−ITD− mutational status, 403 days (95% CI: 259-624) for patients with NPM1+/FLT3−ITD+ or NPM1−/FLT3− mutational status (“Others”) and 342 days (95% CI: 72-NR) for patients with NPM1−/ FLT3−ITD+ mutational status. P-values are estimated with log-rank test.
References


PubMed Central PMCID: PMC4918140.


