Polymorphisms in human homeobox \textit{HLX1} and DNA repair \textit{RAD51} genes increase the risk of therapy-related acute myeloid leukemia

Running title: Target cell size and leukemia risk

Mays Jawad\textsuperscript{1}, Claire Seedhouse\textsuperscript{2}, Nigel Russell\textsuperscript{2} and Mark Plumb\textsuperscript{1}

\textsuperscript{1}Department of Genetics, University of Leicester, University Road, Leicester LE1 7RH, U.K. \textsuperscript{2}Academic Haematology, Clinical Sciences Building, Nottingham University Hospitals, Hucknall Road, Nottingham NG5 1PB, U.K.

CS is supported by the Leukaemia Research Fund, United Kingdom.

Corresponding Author
Claire Seedhouse PhD,
Department of Academic Haematology
Clinical Sciences Building,
Nottingham University Hospitals
Nottingham NG5 1PB
UK
Tel: +44 115 8231823 Fax: +44 115 8231820
claire.seedhouse@nottingham.ac.uk

Mays Jawad designed research, performed research and wrote the manuscript. Claire Seedhouse prepared samples, analysed data and wrote the manuscript. Nigel Russell contributed samples and wrote the manuscript. Mark Plumb designed research and wrote the manuscript

Scientific Heading: Neoplasia Word count: Abstract= 146, Text = 1195
Abstract:

Studies of mouse radiation-induced AML suggest that the number of target stem cells is a risk factor, and the *HLX1* homeobox gene, which is important for hemopoietic development, a candidate gene. The distribution of the C/T-3’UTR polymorphism in *HLX1* in AML and therapy-related AML (t-AML) patients compared to controls was therefore determined. The presence of the variant *HLX1* allele significantly increases the risk of t-AML (OR= 3.36, 95% CI 1.65-6.84). The DNA repair gene *RAD51* (135G/C-5’UTR) polymorphism also increases t-AML risk, and when combined analysis was performed on both *RAD51* and *HLX1* variant alleles, a synergistic 9.5-fold increase (95% CI 2.22-40.64) in the risk of t-AML was observed. We suggest that the *HLX1* polymorphism has an effect on stem cell numbers, whereas an increased DNA repair capacity (*RAD51*) will suppress apoptosis, a genetic interaction that may increase the number of genomes at risk during cancer therapy.
Introduction

Cancer therapy carries up to a 10% risk of a secondary therapy-related acute myeloid leukemia (t-AML)\(^1\). AML is a clonal hematopoietic stem cell malignancy, and low penetrance susceptibility genes within the general population contribute to t-AML risk. The risk of malignant transformation depends on the number of mutations required and the mutation rate, so AML and t-AML genetic association studies have focused on genes encoding proteins involved in maintaining genomic stability\(^2-6\).

One additional theoretical t-AML risk factor is the target cell frequency, as this defines the number of genomes at risk. A mouse genetic linkage analysis of radiation-induced AML (r-AML) identified two r-AML low penetrance susceptibility loci on mouse chromosomes 1 and 6\(^7\). The chromosome 1 susceptibility locus contains the stem cell frequency regulator 1 (Scfr1) locus\(^8,9\), which determines the frequency of mouse bone marrow stem cells. AML-resistant C57 mice have a lower frequency of stem cells than other inbred mouse strains including r-AML susceptible mice\(^7,9\), so stem cell frequency may be part of the equation that defines the risk of r-AML in mice.

The Scfr1 locus (human chromosome 1q41-42) harbors the H2.0-like homeobox (HLXI) gene that is essential for hemopoietic development\(^10\). In humans, high levels of HLXI mRNA are found in CD34\(^+\) bone marrow cells but not granulocytes or macrophages. HLXI mRNA levels are further increased when bone marrow cells are stimulated to proliferate and differentiate in response to cytokines and growth factors, and are particularly high in AML\(^11,12\). Thus, HLXI is implicated in immature stem/progenitor cell biology and, as a homeobox transcription factor, may be involved in establishing the frequency of stem cells during early development.
We have therefore carried out a t-AML patient genetic association study on two *HLX1* gene polymorphisms; one causes an amino acid change (HLX-1:C/T,P365T; NCBI dbSNP: rs2738755) and the other lies in the 3’ untranslated region (HLX1-C/T,3’UTR; NCBI dbSNP: rs2738756). Polymorphisms in the *RAD51* gene promoter (-135G/C,5’UTR), and in the epoxide hydrolase gene (*HYLI*) which is closely linked to *HLX1* on chromosome 1, have both been previously implicated in t-AML risk. As t-AML risk will be highest in individuals who inherit a number of susceptibility genes, genes that encode proteins involved in DNA repair (XRCC2- R118H) and folate metabolism (MTHFR-A667V, MTHFR-E1298A, and MS-A2756T) were also assessed.

**Methods**

**Patient samples**

All blood or bone marrow samples were from Caucasian patients and age-matched controls from the same geographic community and were obtained from Nottingham City Hospital following informed consent as previously reported.

**HLX1 polymorphism analyses**

A 3170bp fragment of the *HLX1* gene (NCBI accession:NM_021958.2) was PCR amplified using forward 5’-CGCTTTAGGTCTTCCGACTG-3’ and reverse 5’-TGCTTCCGGAGAGAAGTGTT-3’ primers. $\gamma^{[32]}$-ATP labeled Allele Specific Oligonucleotides were used to genotype HLX1-P365T (TGGAGACC CGGCTGCGGA and GGAGCCCT GGCTGCGGAT) and HLX1-C/T (3’UTR) (ACTAGGGCGGAGGGGATC and ACTAGGG TGAGGGGATC) by dot blot analysis of duplicate membranes. The *HLX1* genotypes were confirmed by direct sequencing.
Results and Discussion

**HLX1 polymorphisms**

*de novo* AML, t-AML and age-matched control DNA samples were genotyped (Table 1). The variant allele frequencies in the controls were: HLX1-P365T, 0.40; HLX1-C/T (3’UTR), 0.14. Both were in Hardy-Weinberg equilibrium.

The adjusted odds ratios (OR) for each genotype in the AML and t-AML patients compared to controls shows that only the HLX1-C/T (3’UTR) polymorphism achieved statistical significance when the t-AML patients were compared to the controls (Table 1). The proportion of t-AML patients heterozygous (CT) for the HLX1-C/T (3’UTR) polymorphism (45%) was higher than in the control (20%) or *de novo* AML (21%) groups (p< 0.001). The heterozygous HLX1-C/T (3’UTR) CT genotype is associated with a 3.58-fold increase in t-AML risk, with a 3.36-fold increased risk of t-AML in patients who possessed at least one polymorphic T allele. There was no risk when the homozygous variant HLX1-C/T (3’UTR) TT genotype was assessed but this may be attributed to the rarity of this genotype in the general population and the relatively small number of t-AML patients in this study.

The AML patients and controls were assessed for six other functional gene polymorphisms using established methods (XRCC2-R118H, MTHFR-A667V, MTHFR-E1298A, MS-A2756T, HYL1-Y113H and HYL1-H139R)\(^2,6,13-15\). The epoxide hydrolase HYL1-Y113H genotype frequency in the controls was not in Hardy-Weinberg equilibrium (p= 0.002), but was similar to the frequency reported by Lebailly and co-workers (also not in Hardy Weinberg equilibrium)\(^2\). This is probably due to hidden population structures that specifically affect HYL1. The remaining polymorphisms were in Hardy-Weinberg equilibrium (p > 0.14).
Individually, the distributions of the polymorphisms did not exhibit statistically significant differences when comparing the AML patient and controls groups (p > 0.1). In total, the AML samples utilized in this study have been genotyped for 15 polymorphisms in ten genes (RAD51, XRCC1, XRCC2, XRCC3, XPD; NQO1, HYL1, MTHFR, MS and HLXI: this study;5,6). Only the HLX1-C/T (3’UTR) and RAD51 (-135G/C) variant alleles showed any significant association with t-AML risk, and we were unable to confirm that functional polymorphisms in the epoxide hydrolase gene are associated with the risk of AML5. Multiple testing have been performed on these samples, and although there is evidence that applying the Bonferroni corrections may not be appropriate in this sort of genetic association study16, the HLX1-C/T 3’UTR polymorphism (p < 0.001) would still be highly significant if the correction was applied.

To our knowledge, this is the first time that a gene which is implicated in target cell biology has been associated with an increased risk of leukemia. The HLX1-C/T (3’UTR) polymorphism lies in the 3’UTR region of the HLXI gene and further investigations are required to demonstrate that it affects gene function17. Alternatively, it may be associated with either an as yet unidentified functional polymorphism in the HLXI gene, or associated with another nearby causative gene polymorphism. The region between the HLX1-P365T and HLX1-C/T (3’UTR) polymorphisms of the HLXI gene has been sequenced in all the samples in this study and no additional polymorphisms found.

**Combined analysis of polymorphisms in HLXI and RAD51**

The DNA repair RAD51 gene (135G/C variant allele) polymorphism conferred a 2.66-fold (95%CI: 1.17-6.02; p= 0.02) risk of t-AML in these same patients and controls6. The HLX1-C/T (3’UTR) and RAD51-135G/C polymorphisms were therefore analyzed by combined
logistic analyses (Table 2). A strong genetic interaction between the HLX1-C/T (3’UTR) and RAD51-135 G/C was specifically observed in t-AML patients compared to controls, with a significant 9.5-fold increase in the risk of t-AML found in individuals with at least one variant HLX1-C/T (3’UTR) T and at least one variant RAD51-135 G allele. The substantial 9.5 odds ratio for the combined genotype is significantly higher than the sum of the individual OR (3.36 and 2.67), suggesting there is a synergistic rather than additive genetic interaction.

The RAD51 polymorphism leads to enhanced promoter activity and elevated mRNA expression\textsuperscript{18,19}. The increase in cancer risk associated with an increased DNA repair capacity is counter-intuitive, but a highly efficient DNA repair system may suppress apoptosis\textsuperscript{6}. The strong synergistic genetic interaction between HLX1-C/T (3’UTR) and RAD51-135G/C may thus be because they both increase the number of genomes at risk by determining stem cell frequency and by indirectly suppressing target cell apoptosis in response to genotoxic insult.

It must be stressed that the number of t-AML samples assessed in this study was small and much larger cohorts of t-AML samples will be required to confirm the associations described here.

References:


Tables

Table 1
Frequency of HLX1-P365T and HLX1-C/T (3’UTR) polymorphisms in AML and control populations and the relative risk for AML associated with these genotypes. The AML patient groups have been compared with the control group and adjusted for age.

Table 2
Logistic regression analysis on combined genotypes.
Variant (V) genotype includes all heterozygous and variant homozygous genotypes. Wild Type (WT) comprises the homozygous genotype for the most common and frequent genotype. The AML patient groups have been compared with the control group and adjusted for age.
Table 1

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Controls</th>
<th>AML</th>
<th>OR* (95% CI)</th>
<th>( p )</th>
<th>t-AML</th>
<th>OR (95% CI)</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=170 (%)</td>
<td>n=156 (%)</td>
<td></td>
<td></td>
<td>n=41 (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HLX1-P365T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>75 (44)</td>
<td>65 (42)</td>
<td>1.0 (Ref) ▲</td>
<td>-</td>
<td>14 (34)</td>
<td>1.0 (Ref)</td>
<td>-</td>
</tr>
<tr>
<td>CT</td>
<td>76 (45)</td>
<td>76 (49)</td>
<td>1.15 (0.72-1.83)</td>
<td>0.57</td>
<td>22 (54)</td>
<td>1.73 (0.81-3.69)</td>
<td>0.16</td>
</tr>
<tr>
<td>TT</td>
<td>19 (11)</td>
<td>15 (10)</td>
<td>0.84 (0.39-1.82)</td>
<td>0.66</td>
<td>5 (12)</td>
<td>1.02 (0.30-3.51)</td>
<td>0.97</td>
</tr>
<tr>
<td>CT + TT</td>
<td>95 (56)</td>
<td>91 (58)</td>
<td>1.08 (0.69-1.69)</td>
<td>0.73</td>
<td>27 (66)</td>
<td>1.56 (0.75-3.23)</td>
<td>0.23</td>
</tr>
<tr>
<td>( \chi^2, p )</td>
<td>0.75</td>
<td></td>
<td></td>
<td></td>
<td>0.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HLX1-C/T</td>
<td>n=189 (%)</td>
<td>n=166 (%)</td>
<td></td>
<td></td>
<td>n=42 (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3’UTR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>148 (78)</td>
<td>126 (76)</td>
<td>1.0 (Ref)</td>
<td>22 (52)</td>
<td>1.0 (Ref)</td>
<td>1.0 (Ref)</td>
<td>1.0</td>
</tr>
<tr>
<td>CT</td>
<td>37 (20)</td>
<td>35 (21)</td>
<td>1.04 (0.61-1.77)</td>
<td>0.90</td>
<td>19 (45)</td>
<td>3.58 (1.73-7.42)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>TT</td>
<td>4 (2)</td>
<td>5 (3)</td>
<td>1.45 (0.38-5.57)</td>
<td>0.59</td>
<td>1 (2)</td>
<td>1.48 (0.15-14.25)</td>
<td>0.73</td>
</tr>
<tr>
<td>CT + TT</td>
<td>41 (22)</td>
<td>40 (24)</td>
<td>1.08 (0.65-1.79)</td>
<td>0.78</td>
<td>20 (48)</td>
<td>3.36 (1.65-6.84)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>( \chi^2, p )</td>
<td>0.80</td>
<td></td>
<td></td>
<td></td>
<td>0.002</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*OR, odds ratio;

▲CI, confidence interval;

▼Ref, used as reference group.
Table 2

<table>
<thead>
<tr>
<th>RAD51-135G/C (3’UTR)</th>
<th>HLX1-C/T</th>
<th>Control samples</th>
<th>AML samples</th>
<th>OR* (95%CI)‡</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>de novo AML</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT</td>
<td>WT</td>
<td>106</td>
<td>80</td>
<td>1.0 (Ref)‡</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WT</td>
<td>30</td>
<td>22</td>
<td>0.90 (0.47-1.71)</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>V</td>
<td>15</td>
<td>16</td>
<td>1.42 (0.66-3.06)</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>V</td>
<td>4</td>
<td>5</td>
<td>1.80 (0.47-6.97)</td>
<td>0.39</td>
</tr>
<tr>
<td>t-AML</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT</td>
<td>WT</td>
<td>106</td>
<td>17</td>
<td>1.0 (Ref)‡</td>
<td></td>
</tr>
<tr>
<td></td>
<td>V</td>
<td>30</td>
<td>11</td>
<td>2.31 (0.96-5.57)</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>WT</td>
<td>15</td>
<td>3</td>
<td>1.26 (0.32-4.87)</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>V</td>
<td>4</td>
<td>5</td>
<td>9.50 (2.22-40.64)</td>
<td>0.002</td>
</tr>
</tbody>
</table>

*OR, odds ratio;  
‡CI, confidence interval;  
§Ref, used as reference group.
Polymorphisms in human homeobox HLX1 and DNA repair RAD51 genes increase the risk of therapy-related acute myeloid leukemia

Mays Jawad, Claire Helen Seedhouse, Nigel Russell and Mark Plumb