Metabolic Enzyme Polymorphisms and Susceptibility to Acute Leukemia in Adults

Gareth J. Morgan¹ and Martyn T. Smith²
1 Academic Unit of Haematology & Oncology, University of Leeds, Leeds, England
2 Division of Environmental Health Sciences, School of Public Health, University of California, Berkeley, California, USA

Contents
Abstract ......................................................................................................................... 79
1. The Biology and Classification of Acute Leukemia ............................................... 80
2. Descriptive Epidemiology ...................................................................................... 80
3. Genetic Variation in Xenobiotic Metabolism and Leukemia Risk ....................... 81
   3.1 Cytochrome P450 (CYP) .................................................................................. 81
      3.1.1 CYP2D6 .................................................................................................. 81
      3.1.2 CYP2C19 ............................................................................................... 82
      3.1.3 CYP1A1 ................................................................................................. 82
   3.2 Glutathione S-Transferases ............................................................................... 83
   3.3 N-Acetyl Transferases ...................................................................................... 85
4. Insight from Studies with Established Leukemogens .......................................... 86
   4.1 Benzene .......................................................................................................... 86
   4.2 Cytotoxic Chemotherapy ............................................................................... 88
      4.2.1 Alkylation Agents .................................................................................. 88
      4.2.2 Topoisomerase II Inhibitors .................................................................. 89
5. The Folate Pathway and Leukemia ...................................................................... 90
6. Conclusion ............................................................................................................ 90

Abstract

Genetic approaches to understanding the etiology of the acute leukemias are beginning to deliver meaningful insights. Polymorphic variants in xenobiotic metabolizer loci were a natural starting point to study the relevance of these changes. The finding that glutathione S-transferase (GST) T1 null variants increase leukemia risk has implicated oxidative stress in hematopoietic stem cells as an important etiological factor in acute myeloid leukemia (AML). The importance of these enzyme systems in handling specific substrates has also been confirmed by the finding of an increased risk of therapy-related leukemia in individuals with underactive variants of GSTP1 who have been exposed to a chemotherapeutic agent metabolized by this enzyme.

Benzene is a well-recognized leukemogen, and genetic variants in its metabolic pathway can modulate the risk of leukemia following exposure. In particular, underactive variants of the NADP⁺/NADPH oxidoreductase 1 gene (NQO1) seem to increase the risk of AML. Other enzymes within the pathway are proving more difficult to study because of the absence of variants that significantly affect the biological activity of the enzyme under study. No effect of the myeloperoxidase (MPO) gene variants in altering the risk of AML has been seen in our studies.

Another pathway recently shown to be important in determining leukemia risk is folic acid metabolism, particularly important in predisposition to acute lymphocytic leukemia (ALL). Polymorphic variants of the methylenetetrahydrofolate reductase gene (MTHFR) which impair its activity have been shown to be associated with a protective effect. This is thought to be due to an increased availability of nucleotide precursors for incorporation into DNA. This finding implicates misincorporation of uracil into DNA as an important mechanism of leukemic change in lymphoid precursors.

Future studies will extend these observations but will require biological material collected from large well-controlled epidemiological studies. The technological challenges imposed by the high throughput of samples required by these studies are currently being addressed.
Epidemiological data are consistent with a range of susceptibilities to different malignancies distributed throughout the population. Sporadic cancers are likely to be polygenic disorders, in which individual differences in susceptibility depend upon a combination of expression, or function of a range of low penetrance genes. Many of these genes are thought to modulate interactions with the environment – the so-called ‘gene-environment interaction’. Often, in the absence of an exposure, low activity variants exert no deleterious effects and, similarly, at high levels of exposure, no effect upon risk is expected. However, with low to moderate levels of exposure, such as those encountered in the environment, the contribution of host factors governing the response to exposure is likely to be particularly strong.

The completion of the Human Genome Project has focused research activity on the effect of genetic variation within xenobiotic metabolizer genes on susceptibility to disease. Most medically interesting variations in DNA are in the form of single base pair changes, and those occurring at more than 1% of the population are known as single nucleotide polymorphisms (SNPs). SNPs occur one in every 500 base pairs, but as only 3% of the genome encodes useful sequence, it is likely that there will only be around 60 000 medically important variants. The settings in which these genetic variants are usually studied are case-control or cohort studies. Interpretation of the results of these studies is not simple and must be tempered by the need to demonstrate reproducibility of the findings in different populations. In particular, when studying rare diseases such as the leukemias, studies are often small and lack statistical power. They are open to criticism based on this and the many other design issues that affect interpretation of epidemiological research. For the study of genetic variants in this setting it is also important to demonstrate that the polymorphism being investigated exerts a functional effect, that this effect is biologically plausible, and affects a pathway of importance in mediating risk within an appropriate cellular target. Thus, before selecting candidate genes for further analysis, it is important to have a clear understanding of the biology and classification of the disease under study. By so doing, it should be possible to improve the design of studies such that the chances of answering the question addressed can be improved.

1. The Biology and Classification of Acute Leukemia

The normal hematopoietic system is organized as a differentiating hierarchy in which a small number of self-replicating stem cells differentiate to both lymphoid and myeloid lineages. Acute leukemia represents an abnormality of this system, characterized by the accumulation in the bone marrow of blasts, which continue to divide, yet fail to differentiate. It is not a homogeneous disease and can be of lymphoid [acute lymphocytic leukemia (ALL)] or myeloid lineage [acute myeloid leukemia (AML)] and subtypes of each are recognized. AML is related to myelodysplastic syndrome (MDS), a preleukemic condition, and the French-American-British (FAB) classification separates these conditions based on the number of blast cells, the presence of dysplasia, and stage of differentiation of the blast cells. There are a number of recurrent cytogenetic subtypes of AML that have been shown to be of importance clinically, and may also be important in etiologic classification. The largest groups are the balanced translocations, which are present in approximately 20% of all cases of AML, and a group characterized by interstitial deletions of chromosome 5 and 7q, which are present in approximately 16%.[1] The balanced translocations consist of the inv(16), t(15;17) and t(8;21), which have been fully characterized at a molecular level and are associated with good clinical outcome. The 5q/7q- lesion is associated with secondary AML and a poor outcome following treatment.[2]

A further way of looking at etiological groups of AML is based on prior xenobiotic exposure. In this fashion, a group of ‘de novo’ cases with no prior exposure can be defined, together with a group of secondary cases, which have arisen either from MDS or from a prior known exposure. Despite ALL being morphologically more homogeneous than AML, it is also a heterogeneous disease and can be differentiated by clinical behavior and recurrent cytogenetic abnormalities. In adults, the most common cytogenetic abnormality is the t(9;22), which has classically been associated with chronic myeloid leukemia (CML).

2. Descriptive Epidemiology

Classical descriptive epidemiology has shown that AML incidence rates peak slightly in infancy, then decline to age 10, whereupon the incidence begins to rise. After age 40, the incidence rises more rapidly until approximately age 70. During infancy, childhood and early adulthood, incidence rates are similar in males and females. At age 40, AML incidence rises more rapidly in males, with rates subsequently higher in males than in females from middle-age onward.[3] The incidence of the balanced translocations is stable through this age distribution, whereas the incidence of the 5q7q- group increases with increasing age.[4] For ALL, the peak incidence is seen at ages 2 to 4 years, followed by a declining incidence rate throughout the remainder of childhood, adolescence and early adulthood. Rates continue to decline to a nadir at age 40 before rising again with increasing age to reach a peak among the elderly that is lower than the peak in childhood.
Within this pattern, the distribution of cytogenetic subtypes varies, with the t(12;21) being common in childhood disease, and the t(9;22) being more common in adult disease. These descriptive differences argue strongly for the study of etiological factors, not only in the context of lineage but also of cytogenetic subgrouping.

3. Genetic Variation in Xenobiotic Metabolism and Leukemia Risk

A variety of enzymes are critical in the metabolism of a wide range of foreign hydrophobic compounds (xenobiotics), including many potential carcinogens encountered in the environment. The majority of carcinogens require activation to genotoxic electrophilic intermediaries before they can be excreted. The cytochrome P450 (CYP) family is largely responsible for this activity (phase I metabolism), and within these genes there are genetically determined polymorphisms that are associated with the level of in vitro activity of the individual enzymes. The level of activity of these proteins determines the level of genotoxic intermediaries generated, which are further governed by their rate of removal by conjugation and subsequent excretion (phase II metabolism: figure 1). The enzyme systems responsible for phase II activity are also organized into families. The largest and best characterized of these are the glutathione S-transferases (GSTs), N-acetyltransferases, epoxide hydrolases and sulfotransferases, all of which are polymorphic. As the highly reactive products of the phase I enzymes may be carcinogenic, it is clear that expression of phase I and phase II enzymes must be well coordinated. Interactions of ‘at-risk alleles’ within these systems may define people at high risk of malignant disease. Numerous studies have looked at the associations of variants in these genes with the risk of developing malignancy, many of which are summarized in an International Agency for the Research of Cancer (IARC) publication.[14] Despite being good candidates for mediating leukemia risk, the literature in this area is surprisingly small.

3.1 Cytochrome P450 (CYP)

Through the addition of an oxygen singlet, the CYPs activate many environmental procarcinogens to DNA-damaging reactive intermediates. There are multiple family members with different substrate specificity, all with expression located primarily in the liver. Genetic variants of only a few of these have been investigated in leukemia risk, and these are described below.

3.1.1 CYP2D6

CYP2D6 is a non-inducible enzyme whose substrates include components of cigarette smoke, such as 4-(n-methylnitrosamo)- (3-pyridyl)-1-butaneone (NNK) and a variety of therapeutic drugs. The CYP2D6 gene locus on chromosome 22q13.1 encodes more than a dozen CYP2D6 alleles. In Caucasians, a poor metabolizer (PM) phenotype results primarily from inheritance of three mutant alleles, CYP2D6*1, *4 and *5. CYP2D6*4, characterized by a G→A transition at the junction of intron 3 and exon 4, is the most common null CYP2D6 allele, representing almost 75% of mutant alleles. CYP2D6*3 has a single nucleotide deletion in exon 5 and represents 5% of inherited alleles. The CYP2D6*5 allele represents a gene deletion and accounts for 15% of mutant alleles.[5] A significant association between the CYP2D6 PM phenotype and leukemia has been reported in a group of Caucasian patients diagnosed with a range of hematological diseases, including AML and CML.[6] A more recent study contradicted this first study, showing a significant association between the CYP2D6 extensive metabolizer (EM) phenotype and leukemia. The number of patients enrolled in this study was small however, and included a number of different pathological entities.[7] Other studies have reported no association between CYP2D6 polymorphism and hematological disease.[8-10] In a large case-control study, we have shown an association between an increased risk of developing AML and inheritance of the CYP2D6 PM phenotype [odds ratio (OR) = 1.69; 95% confidence interval (CI) = 1.17-2.43; table 1]. A possible role for CYP2D6 in mediating exposure to exogenous factors was also suggested by the analysis of patients with AML as either de novo or secondary.[11] An elevated risk was observed in both groups, but the risk was greater for secondary AML. These data need to be interpreted with caution because of small numbers and because the confidence limits overlap. Interestingly, the CYP2D6 PM phenotype was associated with an increase of AML with a chromosomal abnormality as opposed to AML designated as cytogenetically normal, with
the risk more significant for the 'good prognostic' translocations group than any other classified grouping (Table II). Despite the recognized involvement of CYP2D6 in the metabolism of tobacco components such as nicotine and NNK, no interaction between CYP2D6 phenotype and smoking status was found.

3.1.2 CYP2C19

CYP2C19 is involved in the metabolism of an extensive range of clinical agents, including mephenytoin, omeprazole, and nelfinavir. The CYP2C19 gene locus on chromosome 10q24 is currently known to encode at least eight CYP2C19 alleles. Individuals who are homozygous for the CYP2C19*2 variant allele represent 88% of Caucasian CYP2C19 PMs. The CYP2C19*2 polymorphism is characterized by a G→A transition in exon 5, which results in a truncated and inactive protein. The less common alleles, CYP2C19*4, *5, *6, *7 and *8 are thought to represent the remaining Caucasian PMs. The CYP2C19*3 allele is rarely seen in Caucasians. To date, few studies have investigated the role of CYP2C19 in cancer susceptibility, and only one study in adult acute leukemia has demonstrated an association between the PM phenotype and an increased risk of developing AML (OR = 1.68, 95% CI = 0.97-2.92; Table I). As with studies of CYP2D6, the results also suggested that inheritance of the CYP2C19 PM phenotype was associated with an increased risk of secondary AML, despite the relatively small number of patients involved. A similar pattern of risk was suggested for the CYP2C19 PM as regards cytogenetic classification, with the strongest risk associated with the t(15;17),t(8;21),inv(16) translocation group, but such findings need confirmation in other studies.

3.1.3 CYP1A1

CYP1A1 is primarily an extrahepatic enzyme, expressed in the lung, breast, lymphocyte and placenta. CYP1A1 is highly induced by 3-methylcholanthrene, benzo(a)pyrene and dioxin, all known carcinogens, which could be important leukemogens. Other substrates include the heterocyclic amine, 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine (PhIP), widely found in the diet, and polycyclic aromatic hydrocarbons (PAHs) found in cigarette smoke. The CYP1A1 gene locus located on chromosome 15q22-qter encodes at least five alleles – the wild type CYP1A1*1 allele, and four mutant alleles, CYP1A1*2, *3, *4 and *5. CYP1A1*2 and *4 mutations are located within non-coding regions of the gene. The CYP1A1*3 allele is a newly identified polymorphism and its functional significance is uncertain. CYP1A1*3 is an allele with an A→G point mutation within exon 7 of the gene, a mutation reported to result in a 3-fold increase in the catalytic activity of the enzyme. Because the CYP1A1*3 allele is inherited by only 9% of Caucasians, distribution analysis generally considers CYP1A1*3 heterozygotes and homozy-
Table II. Case-control analysis of de novo adult acute myeloid leukemia (AML) by cytogenetics for CYP2D6 and CYP2C19 status3 (reproduced from Roddam et al.,[11] with permission)

<table>
<thead>
<tr>
<th>Cytogenetics</th>
<th>Case phenotype % (n = 421 total)</th>
<th>Control phenotype % (n = 724 total)</th>
<th>ORq</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EM (n [%])</td>
<td>PM (n [%])</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CYP2D6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any abnormality</td>
<td>166 [86.5]</td>
<td>25 [13.0]</td>
<td>294 [91.3]</td>
<td>24 [7.5]</td>
</tr>
<tr>
<td>translocations</td>
<td>74 [84.1]</td>
<td>14 [15.9]</td>
<td>131 [91.6]</td>
<td>9 [6.3]</td>
</tr>
<tr>
<td>Del/Inv/Trans</td>
<td>18 [81.8]</td>
<td>4 [18.2]</td>
<td>35 [92.1]</td>
<td>2 [5.3]</td>
</tr>
<tr>
<td>other abnormality</td>
<td>74 [90.2]</td>
<td>7 [8.5]</td>
<td>128 [90.8]</td>
<td>13 [9.2]</td>
</tr>
<tr>
<td>CYP2C19</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>143 [95.6]</td>
<td>6 [4.0]</td>
<td>250 [95.4]</td>
<td>12 [4.6]</td>
</tr>
<tr>
<td>translocations</td>
<td>79 [89.8]</td>
<td>9 [10.2]</td>
<td>141 [98.6]</td>
<td>2 [1.4]</td>
</tr>
<tr>
<td>Del/Inv/Trans</td>
<td>20 [90.9]</td>
<td>2 [9.1]</td>
<td>37 [97.4]</td>
<td>1 [2.6]</td>
</tr>
<tr>
<td>other abnormality</td>
<td>79 [96.3]</td>
<td>3 [3.7]</td>
<td>139 [98.6]</td>
<td>2 [1.4]</td>
</tr>
</tbody>
</table>

a Samples are from the Leukaemia Research Fund Adult Acute Leukaemia Case Control Study. Within the above analysis, 7 AML case samples and 8 control samples failed to amplify.
b Cytogenetic classifications used: Normal; translocations including t(15;17), t(8;21) or inv(16); Del/Inv/Trans including -5, 5q-, -7, 7q-, del(12p), -17, 17p-, +8, +11, +21, or abnormalities involving 3q21 or 11q23; Other abnormality includes other clonal abnormalities not classified elsewhere.
c Deduced phenotype
d Odds ratio (OR) estimated using conditional logistic regression, adjusted for deprivation.

CI = confidence interval; Del/Inv/Trans = deletion/inversion/translocation; EM = extensive metabolizers; PM = poor metabolizers.

3.2 Glutathione S-Transferases

The GST gene family may mediate leukemia risk via two potential mechanisms – either by mediating the metabolism of specific leukemogens or by directly affecting the redox potential within the cell, protecting DNA from free radical-induced damage. GSTs play a critical role in the intracellular system which protects against the generation of reactive oxygen species, and in the breakdown of peroxidized lipid and oxidized DNA by the addition of glutathione. GST expression levels can be induced by exposure to foreign substances in vivo,[27] suggesting that they form part of an adaptive system to chemical stress. In addition to this more general role, they are also directly involved in the metabolism of many carcinogens and environmental pollutants. These include environmental carcinogens, such as the toxic epoxides of PAHs, benzo[a]pyrene, acrolein, and other chemicals in cigarette smoke, plus methyl chloride and ethylene oxide. Of particular relevance is the involvement of this family in the metabolism of chemotherapeutic agents that have been implicated in therapy-related AML (tAML).[27,28] In contrast to the CYP gene family, the literature investigating the association of GSTs with leukemia is much more comprehensive.

The gene family includes GST m1 (GSTM1), theta (GSTT1), and pi (GSTP1). Independent gene deletions in GSTT1 or GSTM1 (null genotype) result in a lack of active protein. Two alleles of GSTP1 have been described. GSTP1*B involves an A→G transition, resulting in an Ile→Val substitution in close proximity to the substrate binding site of the molecule, causing an altered specific activity for substrates in comparison with the wild type GSTP1*A allele.[29] The variants described are good candidates for mediating DNA damage in hematopoietic stem and progenitor cells. In vitro, GSTT1 null status has been linked to an increased
Table III. Case-control analysis of adult acute leukemia, acute myeloid leukemia (AML) and acute lymphocytic leukemia (ALL) by GSTM1, GSTT1 and GSTP1 status<sup>a</sup> (reproduced from Rollinson et al.,<sup>b</sup> with permission)

<table>
<thead>
<tr>
<th>GST</th>
<th>Acute leukemia</th>
<th>AML</th>
<th>ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>cases</td>
<td>controls</td>
<td>OR&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(n = 557 total)</td>
<td>(n = 592 total)</td>
<td>no. [%]</td>
</tr>
<tr>
<td>GSTT&lt;sup&gt;c&lt;/sup&gt;</td>
<td>H/H</td>
<td>448 [80.4]</td>
<td>815 [85.6]</td>
</tr>
<tr>
<td></td>
<td>SNA</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>GSTM&lt;sup&gt;c&lt;/sup&gt;</td>
<td>H/H</td>
<td>256 [46.0]</td>
<td>484 [50.8]</td>
</tr>
<tr>
<td></td>
<td>Null</td>
<td>296 [53.1]</td>
<td>466 [49.2]</td>
</tr>
<tr>
<td></td>
<td>SNA</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td><em>A</em>B</td>
<td>204 [36.6]</td>
<td>354 [37.2]</td>
</tr>
<tr>
<td></td>
<td><em>B</em>B</td>
<td>76 [13.6]</td>
<td>133 [14.0]</td>
</tr>
<tr>
<td></td>
<td>SNA</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>GSTP&lt;sup&gt;e&lt;/sup&gt;</td>
<td><em>A</em>A*B</td>
<td>469 [84.2]</td>
<td>813 [85.4]</td>
</tr>
<tr>
<td></td>
<td><em>B</em>B</td>
<td>76 [13.6]</td>
<td>133 [14.0]</td>
</tr>
<tr>
<td></td>
<td>SNA</td>
<td>12</td>
<td>6</td>
</tr>
</tbody>
</table>

<sup>a</sup> Samples from the Leukaemia Research Fund Adult Acute Leukaemia Case Control Study.

<sup>b</sup> Odds ratio (OR) estimated using conditional logistic regression, adjusted for deprivation.

<sup>c</sup> H/H represents individuals who are either homozygous or heterozygous for an active GST allele; Null represents individuals who are homozygous for a deleted GST allele.

<sup>d</sup> *A*A represents individuals who are homozygous for the GSTP1*A allele, *A*B represents individuals who are heterozygous for the GSTP1*A allele and the GSTP1*B allele, *B*B represents individuals who are homozygous for the GSTP1*B allele.

<sup>e</sup> *A*A*B represents combined grouping of individuals with a single or double copy of the GSTP1*A allele. *B*B represents individuals who are homozygous for the GSTP1*B allele.

CI = confidence interval; SNA = sample not amplified.
frequency of diepoxylutane-induced sister chromatid exchange in cultured lymphocytes, while GSTM1 allele status has no effect on the DNA damage observed. Individuals with the GSTP1*B allelic variant have also been shown to have a significantly higher level of hydrophobic DNA adducts. GST T1 and M1 protein and mRNA have been detected within the bone marrow, as has mRNA for GST P1, suggesting that they are important within the target cells for malignant transformation.

There is a significant body of literature that suggests that the GSTT1 null genotype is associated with small but significant increases in the risk of adult acute leukemia. In our study, a significantly higher proportion of patients with AML were found to have the GSTT1 null genotype compared with controls (OR = 1.45; 95% CI = 1.09-1.93) [table III]. As with the CYPs, differences between AML groups with or without chromosomal abnormalities were noted, in particular, the associated risks appeared higher in the group with balanced translocations, t(15;17), t(8;21),inv(16) [table IV], although the significance of this is uncertain. Interestingly, the associated risk seemed higher in the lymphoid cases, despite the fact that this is a much smaller group, and it is also seen in children. However, an association with acute leukemia was not seen in a moderate size US study or a study from Japan looking at de novo and secondary adult AML. Studies of adult MDS, both from the US and Japan, suggest an association, with the latter study showing an even stronger association for the treatment-related form of MDS. Negative studies have also been published for MDS and tAML among children with ALL treated with epipodophyllotoxins. As with GSTT1, a small risk for individuals with the GSTM1 null genotype may exist but again the evidence is conflicting. Ours is the only study that has addressed the role of GSTP1 variants in adult leukemia. This study showed no associated risk for GSTP1 underactive variants and de novo AML. The role of GSTP1 variants in tAML is discussed further in section 4.2.1.

### 3.3 N-Acetyltransferases

NAT2 is a member of the N-acetyl transferase family of enzymes and is involved in the detoxification of aromatic and heterocyclic amines encountered in cooked meats, acetylyating a wide range of compounds into electrophilic nitrenium ions, which are important carcinogens and potential leukemogens. NAT2 is encoded by a highly polymorphic gene for which, to date, 15 allelic variants have been identified. NAT2 activity varies widely, with about 50% of Caucasians demonstrating little or no activity and being classified as slow acetylators. Three main categories of acetylator phenotype have been identified — rapid, slow, and intermediate acetylators — the genotypes of which can
be identified using polymerase chain reaction (PCR). These polymorphic genotypes correspond well to the observed acetylator phenotype as determined using in vivo 'probes', such as caffeine. Both slow [40-46] and rapid acetylator status[47] at NAT2 are considered as risk factors for several malignancies. Rapid acetylators have been associated with increased risks of colon cancer whereas slow acetylators are at increased risk of arylamine-induced bladder cancer.[41] NAT2 substrates are metabolized in two ways, either by N- or O-acetylation. N-acetylation is generally considered to be a detoxifying mechanism, competing with N-hydroxylation mediated via CYP1A2, a reaction known to generate hydroxylamines that can be further metabolized to DNA binding electrophiles.[39] Slow NAT2 acetylation allows substrates to enter the CYP1A2 reaction, yielding a higher level of electrophilic intermediates. The potential predisposing nature of slow acetylator status is suggested by a number of in vivo observations, with increased DNA adduct levels being observed in peripheral blood lymphocytes[48] and some non-hematologic tissues,[149-51] suggesting a possible role of action for NAT2 phenotype in the hematologic system. Although slow acetylator status is a plausible phenotype that may predispose to the development of leukemia, results of the only study looking for an association with adult acute leukemia were negative,[52] even when smoking status was taken into account. The study was large, with adequate power, and therefore, would seem to suggest that xenobiotics metabolized by NAT2 may not be important leukemogens in vivo, even though they may be candidates from in vitro studies. In a study of pediatric ALL, however, a suggestion of an interaction between NAT2 phenotype and other metabolizer enzymes was raised when considered in combination with other genetic variants, including GSTM1.[53]

4. Insight from Studies with Established Leukemogens

4.1 Benzene

The studies described above, looking at supergene families involved in xenobiotic metabolism, exemplify one approach to the investigation of genetically-determined leukemia risk. This is hindered by its failure to take into account known environmental exposures. Another more direct approach is to consider genetic factors that influence susceptibility to established leukemogens. The best described of these associations is with benzene, the metabolism of which has been extensively investigated, allowing us to examine risk based on genetic variants within key metabolic pathways.

Benzene is the oldest and best-known chemical leukemogen, and has been the subject of numerous mechanistic and epidemiologic studies over the years. The first cases of benzene-induced hematotoxicity were described in 1897, and the first case of leukemia associated with benzene exposure was described in 1928. It has been clearly demonstrated that benzene must be metabolized in order for it to induce its hematotoxic and leukemogenic effects. Initial metabolism of benzene takes place in the liver where cytochrome P450, in particular CYP2E1, converts it to a number of reactive intermediates (figure 2). The initial metabolite of benzene is benzene oxide. This compound spontaneously rearranges, mainly to phenol, but may also be conjugated with glutathione by GST to form pre-phenyl mercapturic acid, or metabolized by epoxide hydrolase to benzene dihydrodiol.[54] Benzene dihydrodiol can then undergo dehydrogenation to form catechol, which can be conjugated with sulfate or glucuronic acid and excreted in the urine. Benzene oxide can also form an oxepine, whose ring opens to trans, trans-muconaldehyde, a highly reactive compound that may have toxic effects. The primary metabolite, however, is phenol, which can be further metabolized by CYP2E1 to hydroquinone. Hydroquinone can in turn be hydroxylated to 1,2,4-benzenetriol (figure 2). The three polyphenols, hydroquinone, catechol and benzenetriol, accumulate in the bone marrow and are readily oxidized to highly toxic benzoquinones by peroxidase enzymes, such as bone marrow myeloperoxidase (MPO), a process enhanced by the presence of phenol.

The main protection against the toxic effects of these benzoquinones is NAD(P)H: quinone oxidoreductase 1 (NQO1), originally called DT-diaphorase (figure 2). Thus, people who have low NQO1 activity but high CYP2E1 activity would potentially be more susceptible to the toxic effects of benzene than individuals with low CYP2E1 and high NQO1 activities. A case-control study of benzene-poisoned workers in China[55] showed that this was indeed true, with people lacking NQO1 activity being about 2.5-fold more susceptible to benzene-induced hematotoxicity, and that high versus low CYP2E1 activity conferred an additional 2.5-fold increased susceptibility. People with the combination of high CYP2E1 activity and null NQO1 activity were the most at risk of benzene poisoning (OR = 7.6; 95% CI = 8-31.2).[55] In this study, CYP2E1 activity was measured phenotypically, using hydroxylation of the drug, chlorozoxazone as the biomarker. Unfortunately, this phenotypic measurement does not correlate with any of the multiple genotypes of CYP2E1. It has been suggested that none of the polymorphisms in the CYP2E1 gene have any significant effect on enzyme activity (M. Ingelman-Sundberg, personal communication). Thus, caution should be exercised in studying these polymorphisms in the context of leukemia risk.
NQO1 activity, on the other hand, is closely correlated with genotype. A single nucleotide polymorphism (C→T) at position 609 (C609T) in the NQO1 gene was first identified in a human colon cancer cell line with very low NQO1 activity. This mutation produces a proline to serine substitution that inactivates the enzyme. People who are homozygous for the variant allele completely lack NQO1 activity, while heterozygotes have low intermediate activity compared with the wild-type. The incidence of the polymorphism varies widely by race, and associations have been made between the presence of variant alleles and lung and urological cancers. Given that the NQO1 C609T polymorphism was related to benzene-induced hematotoxicity and leukemia, we investigated whether or not it conferred an increased risk of leukemia in general. Firstly, with Richard Larson’s group in Chicago, we examined the evidence that the NQO1 C609T variant allele was over-represented in tAML, and in those with specific chromosome aberrations in a series of 104 patients with AML or MDS. We found that having low or null NQO1 activity conferred an increased risk of tAML, especially in those with 5q/7q abnormalities. These studies encouraged us to examine the effects of the NQO1 polymorphism in de novo acute leukemias in the general population. DNA samples from a population-based case-control study in England of 493 adult de novo acute leukemia patients and their 838 unaffected age-, sex-, and geographically-matched controls were genotyped for NQO1 C609T. The frequency of cases with low or null NQO1 activity (heterozygous + homozygous mutant) was significantly higher among total acute leukemia cases compared with their matched controls (OR = 1.49; 95% CI = 1.17-1.89; table V). Both ALL (OR = 1.93; 95% CI = 0.96-3.87) and AML cases (OR = 1.47, 95% CI = 1.13-1.90) exhibited a higher ratio of low/null NQO1 genotypes than controls (table V). Among patients with de novo AML, low/null NQO1 activity was significantly associated with the presence of translocations and inversions (OR = 2.39, 95% CI = 1.34-4.27), especially inv(16). These findings were confirmed in a second group of 217 de novo AML cases with known cytogenetics. Our findings were confirmed by a study from Japan in particular the association with secondary AML.

By inference, our data suggests that environmental agents that are normally detoxified by NQO1 are risk factors for producing ALL and AML. This suggests that benzene exposure from gasoline, cigarette smoking and air pollution may be a risk factor for some forms of leukemia. However, it is likely that for most people, environmental benzene exposures are too low to be a significant risk factor. As we have discussed above, NQO1 is thought to protect against benzene by maintaining its phenolic metabolites in their reduced form, preventing quinone formation and reducing oxidative stress. We have recently suggested that
### Table V. Case-control analysis of adult acute leukemia, acute myeloid leukemia (AML) and acute lymphocytic leukemia (ALL) by NAD(P)H: quinone oxidoreductase (NQO1) status

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>NQO1 C609T wild-type (CC)</th>
<th>NQO1 C609T heterozygotes + homozygotes (CT + TT)</th>
<th>OR(^b)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>case [no. (%)]</td>
<td>control [no. (%)]</td>
<td>case [no. (%)]</td>
<td>control [no. (%)]</td>
</tr>
<tr>
<td>Acute leukemia</td>
<td>285 (57.8)</td>
<td>562 (67.1)</td>
<td>205 (41.6)</td>
<td>274 (32.7)</td>
</tr>
<tr>
<td>ALL</td>
<td>36 (55.4)</td>
<td>73 (67.6)</td>
<td>29 (44.6)</td>
<td>35 (32.4)</td>
</tr>
<tr>
<td>ALL-B cell</td>
<td>22 (47.8)</td>
<td>50 (64.1)</td>
<td>24 (52.2)</td>
<td>28 (35.9)</td>
</tr>
<tr>
<td>AML</td>
<td>244 (58.2)</td>
<td>484 (67.3)</td>
<td>175 (41.8)</td>
<td>235 (32.7)</td>
</tr>
<tr>
<td>AML w/trans</td>
<td>46 (52.3)</td>
<td>106 (74.1)</td>
<td>42 (47.7)</td>
<td>37 (25.9)</td>
</tr>
<tr>
<td>t(15;17)</td>
<td>28 (63.6)</td>
<td>51 (68.9)</td>
<td>16 (36.4)</td>
<td>23 (31.1)</td>
</tr>
<tr>
<td>t(8;21)</td>
<td>12 (46.2)</td>
<td>33 (78.6)</td>
<td>14 (53.8)</td>
<td>9 (21.4)</td>
</tr>
<tr>
<td>inv(16)</td>
<td>6 (33.3)</td>
<td>22 (81.5)</td>
<td>12 (66.7)</td>
<td>5 (18.5)</td>
</tr>
<tr>
<td>AML 5q/-q</td>
<td>11 (50.0)</td>
<td>21 (56.8)</td>
<td>11 (50.0)</td>
<td>16 (43.2)</td>
</tr>
</tbody>
</table>

\(a\) Samples are from the Leukaemia Research Fund Adult Acute Leukaemia Case Control Study.

\(b\) Odds ratio (OR) estimated using conditional logistic regression, adjusted for deprivation.

AML w/trans = AMLs harboring common chromosome translocations (t) and inversions (inv); CI = confidence interval.

diet and the intestinal breakdown of excess dietary protein may be a more important source of phenol, hydroquinone and catechol.\(^{57}\) These dietary sources far outweigh environmental benzene exposure, and it is possible that phenols derived mainly from diet are important risk factors for acute leukemia. There are, however, many other compounds that are substrates for NQO1, including quinones, quinone-epoxides, quinone-imines, naphthoquinones, methylene blue, azo and nitro compounds, which may be involved in leukemia induction. Others, potentially metabolized by NQO1, include dietary flavonoids, which are topoisomerase II inhibitors and have been linked with infant leukemia. NQO1 also protects cells from the effects of chronic oxidative stress by maintaining antioxidant forms of ubiquinone and vitamin E. Thus, agents that induce chronic oxidative stress through inflammation or other mechanisms may also play a role in producing acute leukemia.

Figure 2 shows that other enzyme variants may be important in susceptibility to benzene toxicity. For example, the main pathway for removal of the phenolic metabolites from the body is conjugation with sulfate or glucuronic acid. These conjugation reactions are catalyzed by phenol sulfotransferase (PST) and uridine diphosphate glucuronyl transferase (UDPGT), respectively. We are currently investigating the role, if any, that polymorphisms in PST and UDPGT genes play in susceptibility to benzene toxicity and acute leukemia in general. Another key enzyme shown in figure 2 is MPO. Variants of the MPO gene could affect the generation of quinones from phenolic derivatives in the bone marrow. One study has looked at the association of MPO polymorphic variants with the M3 FAB subtype of AML and found a positive association, but this was a very small study of only around 20 patients.\(^{58}\) We have repeated the study on a larger group, including 45 patients with acute promyelocytic (AML-M3) and 62 with myelomonocytic leukemia (AML-M4). In this larger study, we found no difference in variant allele frequency between the cases and their matched controls.

### 4.2 Cytotoxic Chemotherapy

Cytotoxic chemotherapy is the strongest and best described association between exposure and AML. This can potentially be used to model environmentally encountered carcinogens and how variation in their metabolism and repair of the DNA lesions they cause may affect leukemia risk.

#### 4.2.1 Alkylating Agents

Chemotherapy treatment with alkylating agents has been associated with increased risks of MDS and/or AML.\(^{60}\) Typically, they occur five to seven years following treatment, and risk is related to cumulative alkylating drug dose. These conditions are frequently characterized by a preleukemic phase, tri-lineage dysplasia, and cytogenetic abnormalities involving partial deletions of chromosomes 5 and 7. Certain agents, such as melphalan, pose higher risk than others, such as cyclophosphamide.\(^{61,62}\) Therapy-related MDS and AML have been reported following treatment for Hodgkin's disease, non-Hodgkin's lymphoma, multiple myeloma, polycythemia vera, and breast, ovarian, and testicular cancers.\(^{61,63-72}\)

The metabolism of cytotoxic chemotherapy is well understood and GST P1 seems to be particularly important in this respect. Known substrates of GST P1 include melphalan, cyclophosphamide, adriamycin, platinum and busulphan, all widely used chemotherapeutics that have been implicated in the causation of AML. The question of whether the underactive variant of
GSTP1, which carries an isoleucine to valine change at position 105, is associated with an increased risk of tAML has been addressed in a series of 86 patients. This was found to be the case, with a significantly increased risk in patients with exposure to a GST P1 substrate (OR = 4.34; 95% CI = 1.43-13.20). The repair of DNA lesions induced by chemotherapy is also an important pathway for protecting cellular DNA and consequently leukemic transformation. There are also a number of distinct DNA repair pathways that protect against alkylating agent toxicity, including methyl guanine methyl transferase (MGMT), which protects cells from O6 alkyl guanine lesions induced by the chemotherapeutic agents, procarbazine, dacarbazine, carmustine (BCNU), lomustine (CCNU) and temozolomide. MGMT is expressed at low levels in hematopoietic stem cells, and mice deficient for the MGMT gene are markedly sensitive to these agents, suggesting that MGMT exerts important effects in these cells. Recent observations suggest that the mismatch repair (MMR) system can also recognize O6 alkyl guanine lesions and direct cells towards apoptosis. Cells with impaired MGMT and MMR are resistant to apoptosis and have a mutator phenotype. Hematopoietic precursor cells, which have low levels of activity of MGMT and may be unable to fully repair O6 alkyl guanine lesions, would be at particular risk of developing a mutator phenotype if they lost completely or had impaired MMR activity. Recent advances in the laboratory understanding of DNA repair and the description of novel genetic variants that affect the activity of proteins in these pathways are ideal candidates for further testing in association case-control studies.

4.2.2 Topoisomerase II Inhibitors
A second group of tAML is related to therapy with topoisomerase II inhibitors (epidophyllotoxins). In this case, AML is often not preceded by a preleukemic phase, and develops after a shorter latency period of two years. The pattern of cytogenetic abnormalities seen is also different, with balanced translocations involving 11q23 being the most characteristic abnormality, although other balanced translocations are also seen. Variants that affect the metabolism of topoisomerase II inhibitors would, therefore, be ideal candidates for evaluation in case-control studies. CYP3A4 has been investigated in this respect, and has been suggested to be of particular importance.

5. The Folate Pathway and Leukemia
Folic acid metabolism is associated with maintaining DNA integrity and is, therefore, a candidate for mediating leukemia risk. Folic acid is essential in the transfer of methyl groups to various biochemical targets in mammalian tissues involved in amino acid metabolism, and in the synthesis of the purine and pyrimidine components of DNA and RNA. Depleted folic acid levels lead to elevated uracil incorporation into DNA and diminished DNA repair capacity, resulting in DNA strand breaks and chromosomal damage. Impaired folate status modulates the process of neoplastic transformations in selected epithelial tissues (including colon, cervix, lung, and esophagus) and folic acid deficiency has also been associated with neural tube defects in newborns and the elevated plasma homocysteine levels that have been linked with increased cardiovascular disease. We have shown that individuals with specific polymorphisms in the methylenetetrahydrofolate reductase (MTHFR) gene are at reduced risk of adult ALL.

The 5,10-MTHF enzyme, critical in the regulation of folate and methionine metabolism, catalyzes the reduction of 5,10-methyleneTHF (required for purine and pyrimidine synthesis) to 5-methylTHF. A common polymorphism (C677T) in the MTHFR gene results in reduced specific activity of the enzyme, thus affecting folate metabolism. This polymorphism is also implicated in an increased risk of neural tube defects, hyperhomocysteinemia, and occlusive vascular disease. Up to 15% of individuals are homozygous (677T/7) for this allelic variant, with significantly reduced levels of enzyme activity. More recently, a second common polymorphism [glutamate to alanine (A→C) transversion at position 1298 of the MTHFR gene] has also been implicated in neural tube defects. Allele frequencies up to 33% have been previously reported and the variant allele may lead to a reduction in MTHFR enzyme activity.

We have tested whether carriers of variant alleles for MTHFR C677T and/or MTHFR A1298C are protected from adult acute leukemia. We found that the frequency of the MTHFR 677T/T genotype was lower among ALL patients than in matched controls and was associated with a 4.3-fold decrease in risk of ALL (OR = 0.23; 95% CI = 0.86-0.81). Individuals with the MTHFR 1298AC and 1298C/C genotypes were likewise protected from ALL. We observed a 2.8-fold decreased risk of ALL in 1298AC heterozygotes (OR = 0.33; 95% CI = 0.15-0.73), and a 10-fold decreased risk in those with the 1298CC homozygote genotype (OR = 0.07; 95% CI = 0.00-0.63). In AML, no significant difference in MTHFR C677T and A1298C genotype frequencies were observed between cases and controls. Thus, individuals with the MTHFR 677T/T, 1298A/C and 1298C/C genotypes have a decreased risk of adult ALL, but not AML, which suggests that folate metabolism may play a key role in the development of ALL and that misincorporation of uracil into DNA is particularly important in mediating the risk of leukemia transformation in lymphoid progenitor cells.
6. Conclusion

The study of genetic variants in large, well designed case-control studies of acute leukemia has started to yield considerable insights into the role played by specific pathways in governing an individual’s risk of acute leukemia. It seems likely that the genetic approaches outlined in this article will be an area in which significant progress will be made in the next decade. In particular, advances in SNP testing will enable higher throughput analysis, allowing us to examine more variants in larger studies. One such approach that allows rapid screening is DNA pooling, but with more data there will come the need to develop mathematical approaches for effective data analysis. Hopefully, as these approaches are ongoing, biomarkers of exposure will be developed that will enable us to integrate information on exposure and genetic variation to fully explore the role played by gene-environment interactions.

Acknowledgements

We would like to thank P.L. Roddam, S. Rollinson, J.M. Allan and C. Skibola for their original work and contributions to the content of this article. G.J. Morgan is supported by the Leukaemia Research Fund UK. M.T. Smith was supported by NIH grants P30 ES01896, P42 ES04705 and RO1 ES06721 and the National Foundation for Cancer Research. We also wish to thank the MRC Adult Leukaemia Working Party for access to their case series.

References

11. Roddam PL, Rollinson S, Roman E, et al. Metabolisers at the cytochrome P450 2D6 and 2C19 loci are at increased risk of developing adult acute leukaemia. Pharmacogenomics 2000; 10: 605-15
13. Zaphiropoulos PG. RNA molecules containing exons originating from different members of the cytochrome P450 2C gene subfamily (CYP2C) in human epidermis and liver. Nucleic Acids Res 1999; 27: 2585-90


57. McDonald TA, Holland NT, Skibola C, et al. Hypothesis: phenol and hydroquinone derived mainly from diet and gastrointestinal flora activity are causal factors in leukemia [review]. Leukemia 2001; (1): 10-20


73. Glassner BJ, Weeda G, Allan JM, et al. DNA repair methyltransferase (Mgml) knockout mice are sensitive to the lethal effects of chemotherapeutic alkylating agents. Mutagenesis 1999; 14: 339-47


81. Choi SW, Mason JB. Folate and colorectal cancer: is DNA repair the missing link? Am J Gastroenterol 1998; 93: 2013-16

Correspondence and offprints: Professor Gareth J. Morgan, Academic Unit of Haematology & Oncology, University of Leeds, Algernon Firth Building, Leeds, England.
E-mail: garethm@frf.leeds.ac.uk