Pergamon

European Journal of Cancer, Vol. 35, No. 2, pp. 173–185, 1999

Printed in Great Britain
0959-8049/99/$ - see front matter

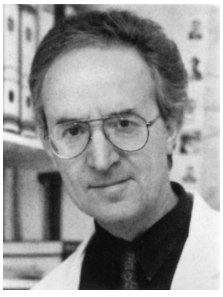
PII: S0959-8049(98)00433-X

European Journal of Cancer
Millennium Review 2000

Molecular Genetics, Natural History and the Demise of Childhood Leukaemia

M. Greaves

LRF Centre for the Cell and Molecular Biology of Leukaemia, Institute of Cancer Research, Chester Beatty Laboratories, 237 Fulham Road, London SW3 6JB, U.K.



After graduating at University College (BSc) and Middlesex Hospital Medical School (Ph.D), Mel worked at the Karolinska Institute, Stockholm, and the National Institute for Medical Research before joining the Imperial Cancer Research Fund, heading up an immunology laboratory in 1976. In 1984, he was appointed Director of the Leukaemia Research Fund's first specialist centre, at the Institute of Cancer Research, London, where he is still based.

The patterns of genetic change, clonal evolution, natural history and latency are very different in the paediatric leukaemias compared with adult epithelial cancers but are similar to those in other childhood cancers of mesenchymal stem cell origin. This distinction has a biological logic in the context of the selective pressures for clonal emergence in different developmental and cellular contexts and has a major impact on curability. Most childhood leukaemias and some other mesenchymal stem cell tumours are of fetal origin and can metastasise without corruption of restraints on cell proliferation or bypassing apoptosis. In marked contrast to most invasive or metastatic epithelial carcinomas in adults, these former cancers then retain sensitivity to therapeutic apoptosis. Moreover, their abbreviated and less complex evolutionary status is associated with less genetic diversity and instability, minimising opportunity for clonal selection for resistance. A minority of leukaemias in children and a higher fraction in adults do, however, have genetic alterations that bypass cell cycle controls and apoptosis imposition. These are the 'bad news' genotypes. The cellular and molecular diversity of acute leukaemia impacts also on aetiology. Paediatric acute leukaemias can be initiated prenatally by illegitimate recombination and fusion gene formation in fetal haemopoiesis. For acute lymphoblastic leukaemia (ALL) in children, twin studies suggest that a secondary postnatal molecular event is also required. This may be promoted by an abnormal or delayed response to common infections. Even for a classic case of a cancer that is intrinsically curable by systematic chemotherapy i.e. childhood ALL, prevention may turn out to be the preferred option. © **1999 Elsevier Science Ltd. All rights reserved.**

Key words: leukaemia, aetiology, differentiation, apoptosis, translocations, chemotherapy, cancer cure, pregnancy, foetus, twins
Eur J Cancer, Vol. 35, No. 2, pp. 173–185, 1999

E-mail: m.greaves@icr.ac.uk
Received 8 Dec. 1998; accepted 9 Dec. 1998.

INTRODUCTION

Good news and bad news

CHILDHOOD LEUKAEMIA has been one of the 20th century's success stories in cancer treatment. This triumph over adversity has been much lauded and deservedly so; for a disease that was, in 1950, universally fatal to have a cure rate of around 75% at the end of the century [1] is no mean feat. Just a few other cancers have succumbed to this challenge—choriocarcinoma, Hodgkin's disease in young adults, Wilms' tumour and testicular carcinoma, all of which are potentially curable by radio- and/or chemotherapy even when metastatic. Sitting uneasily beside this good news is the stubborn refusal of 25% of cases of paediatric leukaemia to go away and, more strikingly, the intransigence of most adult cancers. Only a small minority of the major cancer type—the epithelial carcinomas, are curable once they have breached the point of no return and disseminated. In addition, adult acute leukaemias fare much worse than those in children. I have often asked oncologist colleagues what they believe lies behind this extreme variation in outcome, both between different cancer varieties and within a single type such as acute lymphoblastic leukaemia (ALL). The answer usually goes along the lines of, ''Well, we've known for years that cancers are very diverse in their behaviour, some have cells that are intrinsically sensitive to treatment, others rapidly develop resistance and, oh yes, kids are rather more resilient than old folks as patients.'' Which rather begs the question doesn't it? And the question matters. At the turn of the century, we witness an accelerating race to develop new generation therapeutics for the nasty cancers. This is enabled in equal measure by technological innovation and venture capital mega-dollars and is embellished by media and newsdesk hyperbole. It would be reassuring to say the least if, in setting this agenda, we had more insight into the biological basis of our previous successes and failures. Childhood leukaemia may provide one such opportunity.

At the same time if, as in the past, I ask paediatric oncologists or haematologists a question that parents of children with leukaemia often pose, ''What do you think might cause it and could it be prevented?'' then an apologetic shrug of the shoulders is the commonest response. Which is fair enough; they've not been expected to know, and who does? Their job comes after the event. But the question is again fundamental. Curing a cancer such as ALL can incur a price for a young patient in physical or intellectual development. And even if this was not the case, cancer research driven by the allure of miracle cures is impoverished indeed if it does not pay equal attention to possible causal mechanisms and prospects for prevention. Childhood leukaemia may have something interesting to offer here as well.

What follows then is a biography of a cancer—acute lymphoblastic leukaemia, the commonest subtype of cancer in children. Like all biographies, the perspective is coloured by the author's prejudices but the argument at least is simple: the diversity of the disease can be dissected with the tools of cell and molecular biology. These data can be used to assemble alternative natural histories for ALL which then provide the framework for making sense of what comes before and after a diagnosis: causation and clinical response.

THE BIOLOGICAL DIVERSITY OF ACUTE LEUKAEMIA

The acute leukaemias are broadly divisible by morphological and cytochemical criteria into myeloid and lymphoid varieties. In adults, some 85% are myeloid, the remainder mostly lymphoid. In children the reverse is the case, 80% being lymphoid—ALL. A minority have ambiguous or mixed lympho-myeloid phenotypes. Acute myeloid leukaemias (AML) are clearly heterogeneous in lineage or cell type involvement as revealed by even a cursory scrutiny of morphology but ALL cells have a monotonous appearance that for a long time obscured their origins and diversity.

In the 1970s and early 1980s, it became clear from the phenotypic characterisation with antibodies that ALL was a disease of the B or T lymphocyte lineages, and that the leukaemic lymphoblasts were arrested in their differentiation at what appeared to be equivalent to an early or precursor stage of development [2]. The marked age-associated incidence peak at around 2–5 years for paediatric ALL was found to be exclusively a B lineage disorder [2]. Subsequent analysis of lineage involvement in ALL and other acute leukaemias using chromosomal markers, mutant or rearranged gene markers, immunoglobulin heavy chain (*IGH*) and T-cell receptor (*TCR*) gene rearrangements and X-linked polymorphisms (as clonal markers in females) provided a more stringent test of cellular origins. These analyses substantiated the single cell, monoclonal origin of all leukaemias and provided new insight into the likely developmental level or target cell for mutation and clonal selection [3]. Several acute leukaemias that might have been considered myeloid or lymphoid lineage restricted on the basis of the dominant cellular phenotype turned out to originate in multi-potential lympho-myeloid stem cells. These included a significant proportion of ALL that had a Philadelphia (Ph) chromosome t(9;22)(q34;q11) with either the p210 *BCR-ABL* fusion gene [4] in common with chronic myeloid leukaemia (CML) or the p190 variant of the same gene that is exclusive to ALL [5]. It had been recognised for some time, and before the molecular genetic revolution, that the presence of a Ph chromosome in ALL was a harbinger of bad news as far as clinical outcome was concerned [6]. Furthermore, since considerably more ALL in adults than in children had a Ph^1 chromosome, this provided at least a hint of what might contribute to the age-associated prognosis. Childhood and adult ALL was not the same disease [7]; a subset which responded very poorly in the clinic, originated in a more primitive stem cell and was driven by a unique fusion gene encoding an activated kinase—*BCR-ABL*.

Overall, these clonal analyses provide strong support for the view that whatever the predominant phenotype of the leukaemia cells within a lineage maturation sequence, the molecular lesion responsible for clonal emergence originated in antecedent stem cells. But there are stem cells and stem cells with respect to the hierarchical structure of the haemopoietic system and leukaemias appear to originate at three major developmental levels of lineage specification. In this critical respect, there is a marked bias between the acute leukaemias of paediatric versus adult cases (Figure 1). The likely, though unproven, explanation for this is that lineage committed stem cells are active in early development and provide potential targets for transformation, whereas in adult haemopoiesis predominant stem cells reside at the multipotential lympho-myeloid cell level and hence is preferentially at risk. This designation of cellular origins still leaves unanswered the question of why a group of leukaemias such as adult AML that mostly arise in the same target stem cell population can have such a diversity of phenotypes in

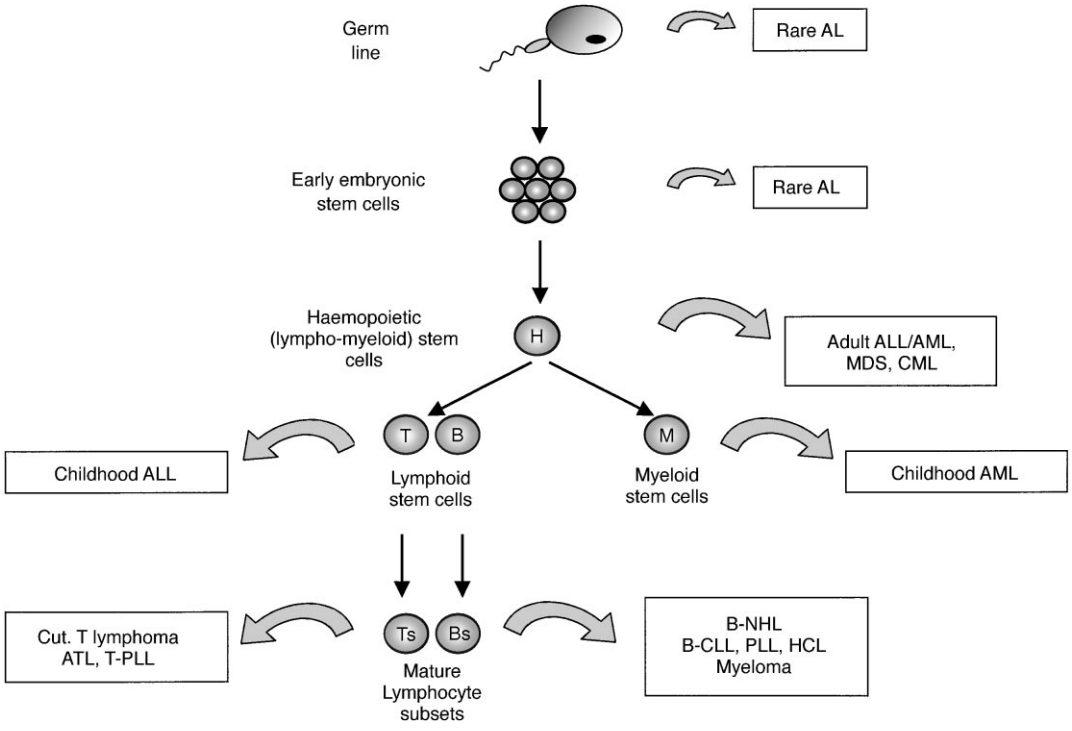


Figure 1. Hierarchical stem cell origins of leukaemia and related cancers. Arrows denote likely level of clonal selection for the majority of cases of leukaemia of subtype listed. AL, acute leukaemia; ALL, acute lymphoblastic leukaemia; AML, acute myeloblastic leukaemia; ATL, adult T cell leukaemia/lymphoma; cut, cutaneous; CLL, chronic lymphocytic leukaemia; HCL, hairy cell leukaemia; NHL, non-Hodgkin's lymphoma; PLL, prolymphocytic leukaemia.

terms of the dominant lineage and apparent level of differentiation achieved. This seems very likely to be a functional attribute of the genes that are driving clonal selection (discussed further below).

Over the past 10 years, a rich tapestry has been assembled of the diversity of molecular abnormalities in acute leukaemia [8]. This now reaches a baroque-like complexity. Over 200 clonal chromosomal changes have been recorded and more than 50 molecular abnormalities identified. As the functions of the proteins encoded by altered genes become uncovered, the extent of diversity becomes less bewildering. Most of the aberrations fall into generic classes of functional dysregulation within which there is considerable redundancy, e.g. cell cycle, apoptosis and differentiation controls [3]. Moreover, a few of the molecular genetic changes are predominant. What stands out from these molecular audits is the abundance of chromosome translocations in which reciprocal chromosomal exchanges of sections of DNA generate novel fusion genes [9, 10]. There are two types (Figure 2). One of these is unique so far to the lymphoid system and involves transposition of a putative oncogene into a hot spot for persistent expression—the *IGH* or *TCR* loci. The other, more widespread type

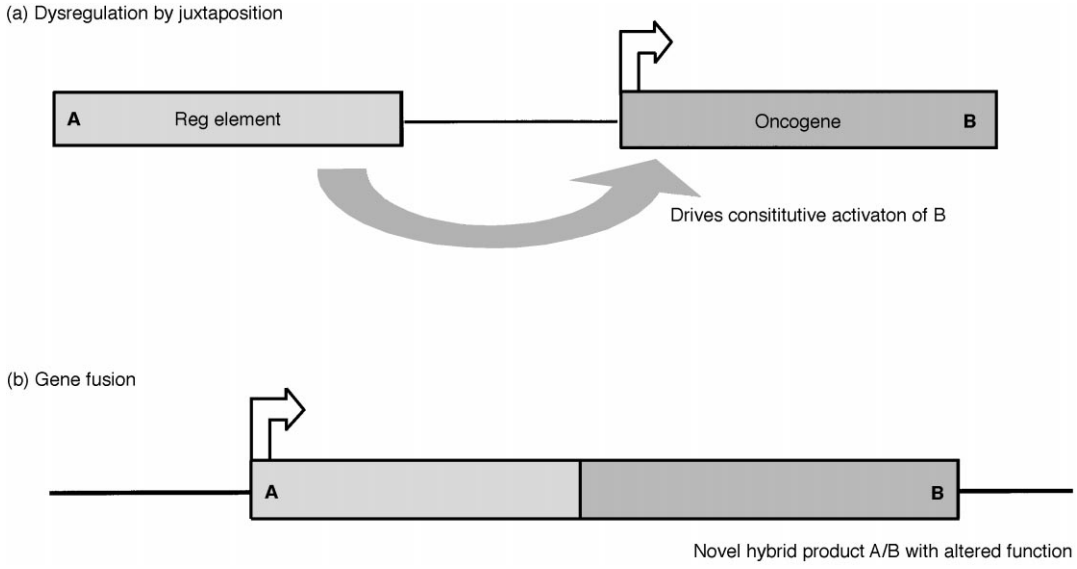


Figure 2. Alternative mechanisms of gene alteration in chromosome translocation. Reg, regulatory element, e.g. *IGH* enhancer.

is actual chimaerism, or fusion, generating a hybrid protein, BCR-ABL being the prototype. These latter dangerous liaisons come in a veritable orgy of combinations. Some individual genes such as *MLL* (in acute leukaemia) [11] or *BCL-6* (in lymphoma) [12] indulge in a plethora of alternative partnerships. Others typified by *TEL* and *AML1* are choreographed into a mix and match dance (Figure 3). In the latter ensemble, there are two major functional outcomes—the constitutive activation of a kinase, or altered transcriptional regulation.

SELECTIVITY OF GENETIC DEFECTS: SPECIFIC STRIKES OR NATURAL SELECTION?

There are several aspects of the selectivity of these chromosomal rearrangements that may carry profound implications for leukaemogenesis and perhaps cancer more generally. The first is the association of particular abnormalities with leukaemia of the same haematological sub-type at different ages. When the relative or proportional frequencies of the major chromosomal abnormalities in ALL are considered in relation to age at diagnosis, then a striking divergence emerges

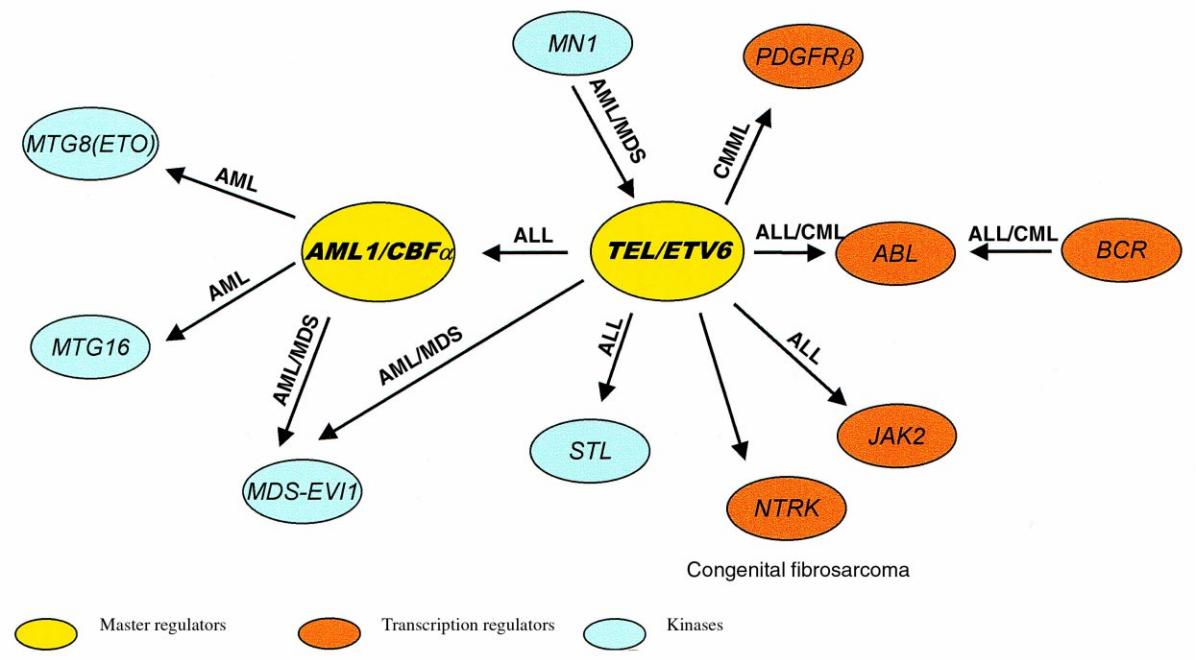


Figure 3. Mix and match gene fusions with *TEL* and *AML1*. Arrows indicate particular gene pairings and associated subtype of leukaemia. ALL, acute lymphoblastic leukaemia; AML, acute myeloblastic leukaemia; CML, chronic myeloid leukaemia; CMML, chronic myelomonocytic leukaemia; MDS, myelodysplastic syndrome (a form of myeloid pre-leukaemia). CBFα, core binding factor α subunit is an alternative name for *AML1*; *ETV6* is alternative name for *TEL*.

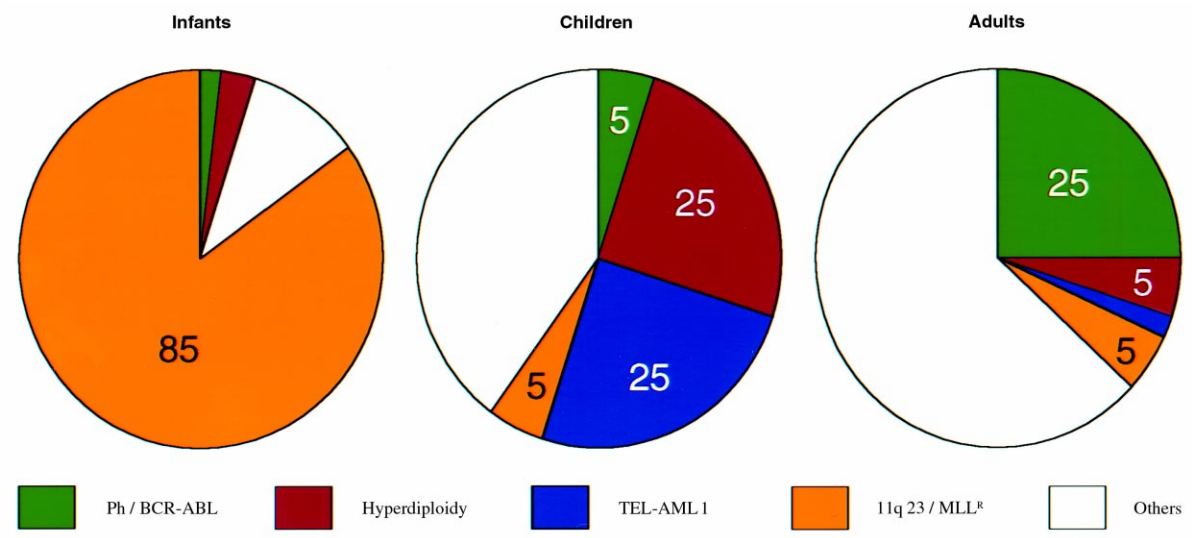


Figure 4. Molecular genetic subsets of ALL. Pie chart shows approximate and relative proportional representation of major molecular abnormalities in cases diagnosed haematologically as ALL in infancy (0–12 months), childhood (1–15 years) and adulthood (>15 years). In reality, these boundaries are less distinct and biological subtypes of ALL overlap in their age distribution.

(Figure 4). In infant ALL (< 12 months), fusions of the *MLL* gene predominate [13]. In the 2–5 year age peak of common (c) ALL, subsets with either hyperdiploidy or *TEL-AML1* fusion are the most prevalent [14, 15]. In adult ALL, the subsets that are frequent in the paediatric cases are very rare and the Ph^1 chromosome, *BCR-ABL* fusion is the most common genetic abnormality, increasing in frequency with age [16].

The second and related aspect of selectivity is the remarkable association between specific fusion gene combinations and leukaemia sub-type as defined by lineage, cell type or level of differentiation (see Figure 3). Janet Rowley (University of Chicago Medical Center, Chicago, U.S.A.) first drew attention to this such highly selective associations in the context of AML and reached the prescient conclusion that the rearrangements provided chromosomal landmarks for the genes that are consistently involved in a cell context dependent manner, possibly linked in some way to differentiation controls [17]. The same is true of ALL and haemopoietic malignancies in general. Chromosomal rearrangements in T lineage ALL subvert different genes to the dictatorial rule of *TCR* gene regulatory elements from those genes similarly coerced by partnership with the *IGH* gene in B lineage malignancies [9]. Within ALL subsets at different ages that all share a predominant early B lineage phenotype, closer scrutiny of the relationship between altered genotype and developmental level of clonal selection reveals further selectivity. Infant ALL with *MLL* gene fusions often have markers indicative of a B/monocytic bilineal origin; cALL with *TEL-AML1* or hyperdiploidy appear to originate in an early B lineage restricted stem cell, whereas as already indicated, *BCR-ABL* fusion often appears to arise in a lympho-myeloid stem cell in adults [18].

These patterns of selectivity could arise by one of several mechanisms [19]. For example, selective transcriptional accessibility in chromatin may provide opportunities for recombination otherwise denied by compact chromatin and gene silence, i.e. a *selective strike*. Alternatively, the rearrangements could occur in a more widespread manner but be registered as leukaemia sub-type associated as a consequence of the selective impact of the protein product of the translocation within a particular cell or developmental context, i.e. by *natural selection*. Both mechanisms could well operate since some degree of physical accessibility is likely to be required. Selective impact of the encoded protein is in fact very likely as indicated by the clonal origin of different subtypes of AML in the same multi-lineage stem cell (see above) and is further supported by experiments in transgenic mice in which fusion genes driven by lineage neutral promoters produce the same restricted cell type of leukaemia in mice as in patients, for example, p190 *BCR-ABL* and B cell precursor ALL [20] and *MLL-AF9* and AML [21]. For ALL with a predominant B cell precursor leukaemic cell phenotype, it can be concluded that the *TEL-AML1* fusion gene and hyperdiploidy could arise by whatever mechanism in a variety of cell types, but that there is selective impact exclusively on a B lineage restricted stem cell that is most active early in life. Alternatively, *BCR-ABL p190* kinase activation may produce a similar leukaemic cell phenotype and level of differentiation block in ALL, but endow clonal advantage in a more primitive lympho-myeloid stem cell that is more at risk in adults.

There is a third striking aspect of selectivity of these genetic recombinations. The generation of fusion genes by chromosomal translocation is a predominant mechanism of cell transformation in stem cells of mesenchymal origin, the leukaemias (plus lymphomas and myelomas), rhabdomyosarcomas, Wilms' tumours and other sarcomas [9, 19]. In contrast, to date, the same genetic gymnastics have made a very timorous appearance on stage for the major adult cancers—the epithelial carcinomas. The explanation for this deficit is either trivial or profound. Owing to the technical difficulties of obtaining good metaphase chromosome spreads of epithelial cancer cells, it could be that so far we have simply missed many underlying chromosomal translocations. If so, then this situation will soon be resolved with improved cell sorting and cell culture methods combined with the application of newer pan-chromosome fluorochrome-labelling methods. Alternatively, and as I suspect, the apparent deficiency of consistent reciprocal translocation may be real and, therefore, much more interesting, particularly if it turns out that dysregulation of transcriptional controls by fusion gene formation is restricted to mesenchymal cancers. If so, then there are several potential explanations [22]. A selective strike at loci accessible in a certain class of stem cells is one possibility, but the more likely option is that this kind of genetic abnormality, as an initiating event, leads primarily to inhibition of differentiation which provides clonal advantage to mesenchymal-derived clones but less so, if at all, to epithelial stem cell derivatives. More on this in a moment.

Clearly the selectivity of molecular abnormalities in leukaemia still awaits a full molecular explanation but it certainly suggests that the developmental and cellular context of a molecular aberration is critical to its likely impact. These observations also reveal that ALL is both biologically and clinically a very different disease in infants, children and adults. This has important implications for treatment, prognosis and possible cure, as we will see later. It is also likely that age-associated subtypes are distinct aetiologically, which is also no trivial matter.

A NATURAL HISTORY TALE

These molecular associations are presumed to be causal, integral components of the pathway of aetiology and pathogenesis. But exactly where do they fit into the natural history of ALL? Are they initiating events and, if so, when do they occur in the prior life history of a child diagnosed, say, at the average age of 3 years? And if they are initiating events, what else is required for evolution of the leukaemic clone to expand sufficiently to prompt diagnostic symptoms? How many accumulative genetic events does it take to get ALL? One thing is clear, children diagnosed with ALL usually have a very short clinical history. Four to six weeks before symptoms of paleness, lethargy, bruising or infection prompt a diagnosis, the patient will usually have been perfectly healthy. Hence the sense of shock and disbelief in parents. Since this cancer is unlikely to have a latency of a few weeks, we can assume that there must be a more protracted but covert and clinically silent natural history. This has now been teased out of cover.

The common molecular abnormalities are acquired or non-constitutive; they are absent from non-blood cells and disappear during remission. For 95% of ALL and AML, there is no evidence for predisposition via the inheritance of dominantly active or highly penetrant, mutant genes. The rare exceptions are cases associated with syndromes associated with constitutive defects in DNA repair—Bloom's syndrome, Fanconi's anaemia, ataxia telangiectasia (AT). For

the latter situations, the risk of acute leukaemia is greatly increased but the cases are usually AML or, in the case of AT, T cell leukaemia [23]. Down's syndrome children have an overall 20–30 times extra risk of acute leukaemia [24] and, for reasons that are still poorly understood, with Down's it is often of a normally very rare sub-type—acute megakaryoblastic. Family pedigrees with acute leukaemia do occur but are rare. There is, therefore, little or no evidence that the predominant sub-types of ALL in infants, children and adults are associated with inherited genetic abnormalities, despite the very young age of many of the former patients. This is not to say that ALL and AML are not subject to genetic influence—in, for example, risk factors linked to causal pathways [23].

If the common fusion genes—say *MLL-AF4* in infant ALL and *TEL-AML1* in childhood ALL (Figure 5), are indeed initiating events, then their origin must be constrained to quite a brief developmental period between the beginning of haemopoiesis or lymphopoiesis in the embryo or fetus (fetal liver at 9/10 weeks and bone marrow at 18–20 weeks) and some time-point, a few months perhaps, prior to diagnosis. Given this narrow window, one might anticipate that the initiating mutations commonly occur prenatally, particularly for infant ALL. The average age for these patients at diagnosis is 6 months but a minority are diagnosed congenitally or neonatally with *MLL* fusion genes [25, 26]. A case of fetal death due to an acute leukaemia with a rearranged *MLL* gene has also been described [27].

Epidemiological studies provide some indirect evidence for fetal origin of paediatric acute leukaemia, though this is more marked for infant AML than for infant or childhood ALL [28]. Mathematical modelling also provides support for the view that a fetal initiation is plausible for typical childhood ALL (i.e. cALL at age 2–5 years) [29]. We have sought more direct, molecular evidence that the common fusion genes in paediatric ALL are initiating, fetal events, and, collectively, the evidence is now persuasive.

IDENTICAL TWINS WITH CONCORDANT ALL

Over 100 years ago, Francis Galton, cousin of Charles Darwin, introduced the idea of comparing identical (monozygotic) with non-identical or fraternal (dizygotic) twins as a means of comparing the relative importance of nature (or inherited genetics) versus nurture (or environment) in normal and pathological features. Most human diseases have now been subject to the twin test, as have most other human features or traits including cognitive ability, self-esteem, sexual preference and obesity. When twins share the same disease or feature, this is described as concordance. Concordance is compatible with a strong genetic influence but does not provide definitive proof (in the absence of knowledge of the genes involved) and could be confounded by the likelihood that identical twins may be more likely than fraternal twins to share a near identical environment. This potential distinction extends to the period before birth when some 60% of monozygotic twins share a single or monochorionic placenta [30]. One hundred per cent of dizygotic twins have independent or dichorionic placentas. This difference turns out to matter with respect to leukaemia.

The concordance rates for most diseases studied reflect an involvement of both inherited predisposition and postnatal environmental exposures (and chance). Rates for autoimmune diseases such as insulin-dependent diabetes and multiple sclerosis are between 25 and 50% [31, 32]. In leukaemia and related diseases, the rates depend critically upon sub-type (Table 1). The high concordance rates for acute

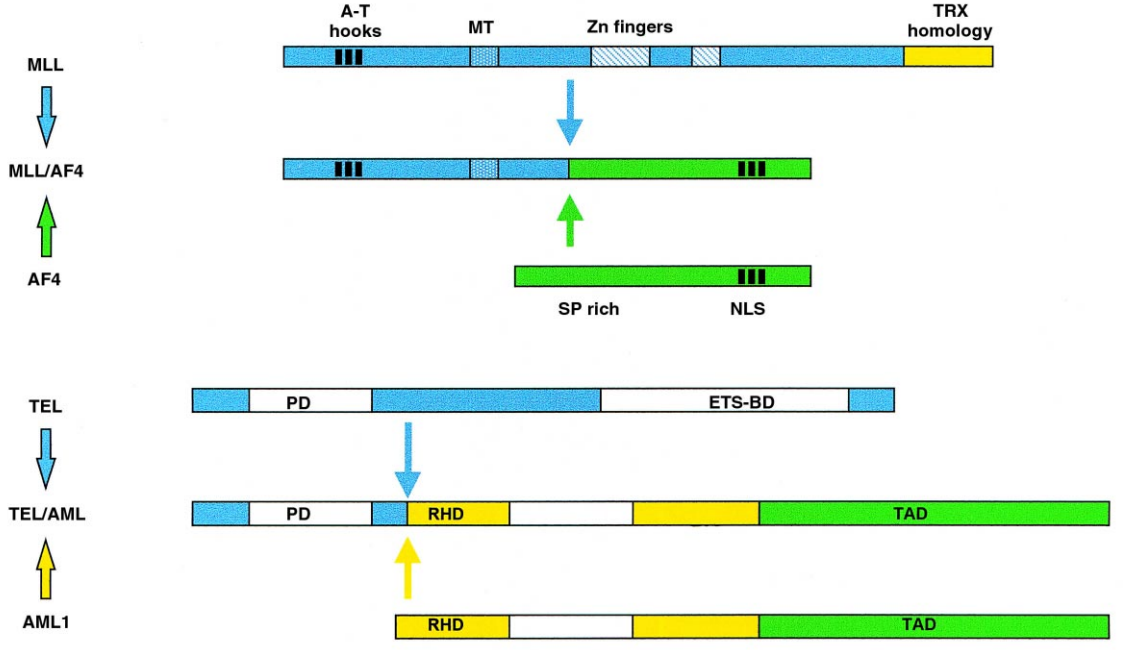


Figure 5. Structural features of the common fusion genes in paediatric acute lymphoblastic leukaemia. Simplified functional domain structure of rearranged, hybrid proteins. AT hooks, adenine, thymine rich region (DNA binding?); Zn, zinc fingers—DNA and/or protein binding; TRX homology-region with sequence homology to the *Drosophila* developmental protein trithorax; SP, serine proline rich—DNA binding?; NLS, nuclear localisation signal. PD, pointed domain (for protein interaction); ETS-BD Ets-like binding domain (to DNA); RHD, Runt Homology Domain—region with sequence homology to Runt gene product in *Drosophila* (DNA binding); TAD—DNA transactivating domain.

leukaemia in the very young have been recognised for many years, the first cases being described in the 1930s. In the 1970s, there was a flurry of interest in such twin pairs and much speculation on its significance [33, 34]. The clearest and most innovative suggestion came from Boyse and Clarkson at Memorial Sloane Kettering, New York, U.S.A. [35]. Boyse as an immunologist was well informed about blood cell chimaerism in twin cattle; this had led Peter Medawar to the discovery of immunological tolerance for which he won his Nobel Prize in 1960. The simple idea was that concordance reflected metastasis from one twin to the other, the leukaemia arising in one fetus with clonal progeny disseminating to the other co-twin via the vascular anastomoses that occur within a single, monochorionic placenta [35].

There is a simple test that discriminates between the Boyse and Clarkson explanation of twin concordance and the 'conventional' explanation for shared illness in identical twins. If the former is correct, then the leukaemic cell population in a pair of twins will belong to a single clone as opposed to two independent clones. Of course the twins are themselves, in genetic terms, members of a single unique clone, but molecular markers that detect post-zygotic acquired genetic variability can distinguish singular from independent clonal origin of cancer cells, even in identical twins.

Several molecular markers of clonality are available [2] but the most decisive and relevant here are the precise breakpoints or fusion sequences in the chimaeric genes that are consistently generated, in ALL, by rearrangement and fusion (Figure 5). DNA breakpoints are clustered in a limited region of noncoding introns (usually several kb) but each patient's leukaemic cells' breakpoint is unique or clonotypic. This genomic variation is usually of no functional significance as subsequent exon splicing ensures that the same chimaeric protein is assembled. This unique feature can be revealed by Southern blot analysis and genomic sequencing of the fusion gene. Using this approach, we found that in three pairs of identical infant twins with *MLL* gene rearrangement and ALL, each pair shared the same clonotypic *MLL* gene break (Figure 6) [36]. Two other pairs of infants with concordant ALL or AML have subsequently been reported to have identical *MLL* gene fusions [37, 38]. Essentially the same result has been now obtained with older children with ALL also. In a pair of Brazilian twin children diagnosed at the age of 9 years with T-ALL in one case and the age of 11 years with T-non-Hodgkin's lymphoma (NHL) in the co-twin, the leukaemic cells shared a common clonal origin as indicated by their unique T-cell receptor β DNJ sequence [39]. More recently, several twins with common ALL and *TEL-AML1* gene fusion have been examined for clonality. This was more technically demanding as the *AML1* breakpoint region is very large (possibly ~100 kb) and no complete genomic sequence was available. To circumvent these difficulties, the *TEL-AML1* genomic fusion region was either directly cloned from a 12 kb Southern blot rearranged restriction fragment identified with a *TEL* probe [40] or amplified and directly sequenced using an adaptation of a long distance inverse PCR method [41]. In three pairs of twins, it was again found that the twins shared the same clonotypic fusion gene sequence and must, therefore, have had a single cell origin in one fetus [40–42]. One of these twin pairs was especially informative. One twin was diagnosed at the age of 5 years, her twin sister 9 years later at the age of 14 years [41]. This divergence of diagnosis in time indicates that the postnatal latency in ALL following prenatal initiation can be variable and occasionally very protracted. At the time twin 1 was diagnosed at the age of 5 years, a routine bone marrow aspirate was taken on her healthy sister. This showed no morphological signs of leukaemia and was archived. Armed with the genomic *TEL-AML1* fusion sequence of the leukaemic clone that emerged in this twin some 9 years later, we could show, using conventional polymerase chain reaction (PCR), that the leukaemic or 'pre-leukaemic' clone was present in the 'normal' bone marrow almost a decade before its emergence to clinical distinction [41]. Furthermore, since the routine bone marrow aspirate is essentially a random, single sampling of marrow, we can conclude from this observation that the 'pre-leukaemic' fetal clone was widely disseminated but subject to some kind of homeostatic or self-limiting control in its expansion.

Table 1. *Concordance rates for haematological malignancies*

Disease subtype	Concordance rate*%
Adult (> 15 years) acute leukaemias†	< 1
Hodgkin's disease in young adults (15–45 years) [84]	5
ALL in children‡ (1–15 years)	~5
Infant (< 12 months) acute leukaemia‡	25–50

*Very approximate. †Limited data available; could be higher in B chronic lymphocytic leukaemia which can be familial. ‡Estimates [18, 85, 86].

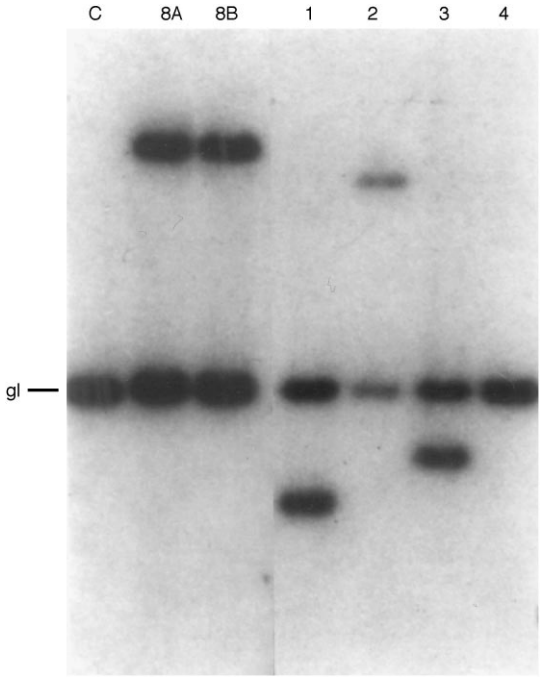


Figure 6. Monoclonal origin of *MLL* gene rearrangement in identical twin infants with ALL. Southern blot showing restriction digest products of the *MLL* gene. 1, 2, 3, DNA from three individual unrelated infants with ALL; the three rearrangements are different. 4, DNA from leukaemic cells (ALL) without *MLL* gene rearrangement. C, control DNA; germ line (gl)/unrearranged *MLL* gene. 8A/B, DNA from leukaemic cells of twin pair showing same rearrangement (also observed with other enzyme restriction digests. No rearrangement seen with constitutive DNA). For further details of the molecular analysis of these twin cases see [36].

The sharing of these clonotypic molecular markers can only be plausibly explained by a prenatal origin in one haemopoietic stem cell with the fusion gene formation representing a very early and likely initiating molecular event. Progeny of this transformed cell then will have metastasised to the co-twin via intraplacental anastomoses as predicted by Boyse and Clarkson. Since monozygotic twins with ALL are in all respects, other than having been accidentally cloned, no different from other paediatric patients with leukaemia, it seems very likely that ALL can have a fetal origin in non-twinned paediatric patients; and maybe in most cases

BACKTRACKING LEUKAEMIA TO BIRTH

This prenatal interpretation of the twin data has now been substantiated in both twins and singletons by a direct molecular test. If pre-leukaemic cells originate in the fetus and, in the identical twin context, spread from one twin to the other before birth via blood, then there is a reasonable likelihood that they will be present in blood at birth in both twins and singletons. All children in most European countries and the U.S.A. have a small blood sample taken by heel prick in the first year of life—the Guthrie card (Figure 7). These neonatal blood spots have been traditionally used for screening for inborn errors of metabolism (phenylketonuria, PKU). It is known, however, that DNA remains intact on old filters and can be amplified by PCR to reveal both constitutive mutations in, for example, the β haemoglobin gene [43] or the CF gene [44] and exogenous viral sequences [45, 46]. We reasoned that it might be possible to detect low copy number leukaemia fusion gene sequences in such material using PCR with patient or clone specific primers. Artificial blood spots in which normal blood was spiked with leukaemic cells indicated that under optimal conditions, 1–10 copies of the genomic fusion sequence (or cells) could be detected. Guthrie cards were first obtained from 3 patients diagnosed with *MLL-AF4* fusion gene positive ALL aged 5, 6 and 24 months. All 3 gave positive PCR signals in single segments excised from their corresponding blood spots [47]. Sequencing of the products confirmed that they were identical to those in the corresponding leukaemic cell DNA at diagnosis. More recently Guthrie cards from newborns who subsequently developed cALL with a *TEL-AML1* fusion gene have been collected. As proof of principle, blood spots from a pair of identical twins, both aged 3 years at diagnosis, were tested [42]. In each case, blood spot slices were positive for the clonotypic sequence, suggesting the presence of a small number of 'pre-leukaemic' cells in blood at birth. Subsequently, a Guthrie card from a singleton with cALL (aged 2 years 1 month) was assessed and was also positive. Other neonatal blood spots from individuals who have later developed *TEL-AML1* positive ALL are currently being assessed. These data provide direct evidence for a fetal origin of

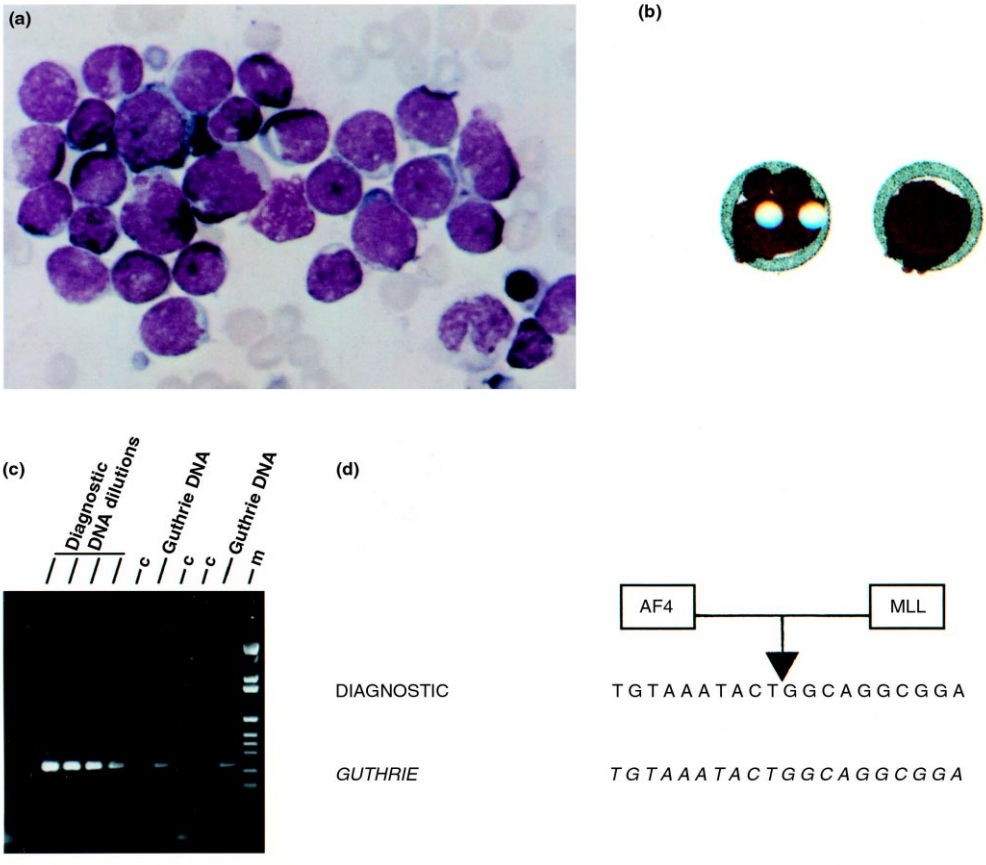


Figure 7. Detection of clonotypic *MLL-AF4* fusion sequences in neonatal blood spots. (a) Infant ALL with *MLL-AF4* fusion gene at diagnosis (source of DNA)—stained bone marrow smear. (b) Guthrie card blood spots. (c) PCR result (stained gel) using clonotypic, patient-specific primers to assess sensitivity of detection of diagnostic *MLL-AF4* DNA, by dilution and detection of *MLL-AF4* fusion in individual (1/10th) slices of blood spot. c, control, slices from other Guthrie cards; m, molecular weight markers. (d) Identical fusion region sequences identified in leukaemic/diagnostic and neonatal blood spot DNA. See [47] for details.

leukaemia fusion gene, both for *MLL-AF4* in infants and *TEL-AML1* in older children. It will almost certainly not be possible to use this approach to resolve whether or not all paediatric cALL are fetal in origin, though most may be. Guthrie blood spots carry a very small volume of blood (~30 μl) and the absolute number of leukaemic/pre-leukaemic cells is likely to be small. Some Guthrie cards will probably yield negative results but such data will not be interpretable.

POST-NATAL GENETIC EVENTS

The concordance rates for leukaemia in identical twins will reflect the requirements for additional events after the initiation of leukaemogenesis in the fetus. For infant ALL with *MLL* fusion genes, the concordance rate is very high. Accurate figures are not available but rates are around 25–50% and may approach 100% for those with a single placenta. Latency is also remarkably brief. As it is in one other situation involving *MLL* gene fusion. This is in so-called secondary leukaemias (usually AML) that arise in a small minority of cancer patients as a consequence of the therapeutic use of genotoxic drugs—as collateral damage [48]. In cases associated with prior use of anthracyclines or epidophyllotoxins (VP16, VP26), rearrangements of the *MLL* gene are common and latencies are very brief, averaging around 18 months. An implication of the very high concordance rate plus very brief latency or time frame of clonal evolution to clinical disease is that the *MLL* gene fusion may be *sufficient* for leukaemogenesis. Alternatively, the fusion gene would have to provoke very efficiently whatever other secondary changes are necessary.

For cALL in children, the story is different. The precise concordance rate for identical twins is again not known but is around 5% with an average postnatal latency of 3 years [18]. This then implies that fetal initiation, by *TEL-AML1* is *insufficient* for overt leukaemia development. From which follows two important predictions. Firstly, that many more children are born with a *TEL-AML1*-driven pre-leukaemic clone than ever become diagnosed with leukaemia. This is currently being assessed by molecular screening of unselected cord blood samples. Secondly, that something else is required after birth for cALL to develop. As with other twin studies of disease, the suspicion then falls on environmental exposures that, in this case, might provoke or promote evolution of the 'pre-leukaemic' clone. In molecular terms, the only consistent genetic abnormality observed at diagnosis in cALL along with *TEL-AML1* fusion gene itself is deletion of the normal *TEL* allele from the parental chromosome 12 not involved in the translocation. This occurs in most cases of cALL with *TEL-AML1* fusion [49] and fluorescence *in situ* hybridisation (FISH) studies provide evidence that *TEL* allele deletion is secondary or sub-clonal to *TEL-AML1* fusion [50]. This combination of genetic abnormalities is beginning to make sense with respect to the likely biochemical function of the chimaeric fusion protein. *TEL (ETV6)* and *AML1 (CBFα)* are essential or obligatory transcriptional regulators in fetal haemopoiesis as revealed by homologous recombination knockouts in mice [51, 52]. The chimaeric protein has altered DNA interactive capacity (Figure 6) compared with either wild type proteins, and transfection experiments suggest that it may function as a dominant negative suppressor of *AML1* target genes, perhaps blocking differentiation [53]. *TEL* has a homodimerisation domain and as a consequence, the presence of a normal TEL protein in a cell with a *TEL-AML1* fusion is likely to result in TEL to *TEL-AML1* heterodimerisation [54] and compromise of fusion protein activity. There is, therefore, advantage to be gained by loss of the normal *TEL* allele and this could represent *a* or *the* critical postnatal event.

WHAT IS PULLING THE STRINGS?

The scenario painted provides the time frame for a brief natural history but is missing a vital ingredient: what is really determining whether or not it starts at all and whether it finishes up as ALL? These genetic events may represent the proximal causal events in ALL but what causes them? In one sense, this aetiological question is a matter for epidemiology but I argue that you would have to strike very lucky indeed to uncover key exposures for leukaemia in the absence of any insight into the biological diversity and natural history of the disease. There are clues, and false trails. The epidemiological evidence is highly suggestive that paediatric leukaemia and cALL in particular is more prevalent in more socio-economically advanced societies. In this context there is an inventory of potentially offensive exposures that is both intriguing and a challenge to credulity. This includes pesticides, electro-magnetic fields, ionising radiation, vitamin K, hot dogs and hamburgers. In principle, if genotoxic damage is the route to leukaemogenesis, then multiple *different* exposures could be involved and the demonstrable strength of association of any one of them would be weak. There is also a difficulty here with respect to epidemiological 'lumping' of disease types. The biological and clinical diversity of leukaemia in infants, children and adults is such that it would be very surprising if they shared a single causal mechanism and more recent epidemiological evidence indicates that they are indeed aetiologically distinct.

With respect to cALL that will include both the subset with *TEL-AML1* fusion plus those with hyperdiploidy, evidence is accumulating that supports an infectious mechanism. The data are reviewed elsewhere [55–57] but in essence, it involves studies on the sociodemography of the disease including population density and population mixing, social contacts, clustering and community socio-economic status. Collectively these data provide a strong hint that some kind of abnormal response to infection is involved. In these circumstances, it is tempting to play the animal precedent card and to implicate leukaemogenic viruses and defective immunological control. This parsimonious explanation is difficult to rule out definitively but it probably is not the mechanism. For two reasons: first, there is no increased risk of cALL in immunosuppressed individuals (in contrast to EBV-associated lymphomas); second, attempts to identify candidate viral sequences in cALL cells have so far failed [58] (R. Jarrett, J. McKenzie and M. Greaves, Institute of Cancer Research, London, U.K.). An alternative explanation accords with the model of the natural history of cALL derived from the experiments described above (Figure 8). What is required postnatally is some exposure or event that can promote the conversion of the fetal pre-leukaemic clone into a fully qualified leukaemia, probably via loss of the second *TEL* allele. Classical *promotion* via proliferative stress would fulfil this role. Rather than failure of the immune response to deal with a virus being the problem, I have suggested that it's the opposite, an over exuberant immune response providing the requisite insult [57]. There are a number of ways in which this might operate including via transient marrow hypoplasia

induced by T-cell derived γ interferon followed by florid regeneration. There are both animal [59] and clinical [60] precedents for lymphoid malignancies originating via indirect immunological or infectious mechanisms. The nature of the infections involved remains entirely conjectural at present, though there is no logic in considering only viruses; bacteria, and particularly common bacteria with super-antigens, are plausible candidates. The author's preference here is, therefore, for a scenario that has an historical resonance.

> We incline on our evidence to the belief that the solution of the problem of leukaemia lies rather in some peculiar reaction to infection than in the existence of some specific infective agent. F.J. Poynton, H. Thursfield and D. Paterson, Great Ormond Street Hospital for Sick Children, 1922 [61].

The extra ingredient of the current hypothesis, other than its second stage promotional role, is that the immune response is predicted to be unbalanced because of a failure (in more 'developed' societies) of appropriate immune-modulating infectious exposures in infancy, plus a possible role of HLA or other immune response genes in influencing patterns of immunological response [57].

These ideas are being assessed in current case/control epidemiological studies in the U.K., U.S.A. and elsewhere. Several predictions follow from the model. Relative risk for cALL is anticipated to decrease with increased social contacts (via siblings or in playgroups for example) in infancy (more opportunities for early infection), with increased recorded common infections in infancy and with protracted breast feeding (modulation of the infant immune response). There is now evidence supporting all three of these predictions [62–64] though the more stringent epidemiological assessment is still to come—in the U.K. National Case–Control Study, data from which should be available by early 2000.

Even if the postulated delayed infection does turn out to have a key promoting role in paediatric ALL, we are left with the conundrum of what causes the initiating *TEL-AML1* recombination event in fetal haemopoiesis. A very neat and convenient explanation would involve misdirection of the RAG1/2 recombinase machinery via *IGH*-like signal sequences [9, 65]. But, to plagiarise Mark Twain—another beautiful hypothesis is ruined by an ugly fact: the required consensus sequences are just not there, or at least not consistently. Neither are there any obvious sequences to facilitate homologous recombination [66]. The same applies to the *EWS-FLI-1* fusions in Ewing's sarcoma [67]. This, in a sense, leaves us in the dark. But the mechanism could be misdirection of conventional DNA PK-mediated double strand break repair, the initial breakage reflecting oxidative damage. The latter could arise from genotoxic exposures (for which there is currently no epidemiological evidence) but equally could be an occasional consequence of the very active proliferation and oxidative metabolism of lymphoid progenitor cells in early development, i.e. a developmental accident.

I draw two main conclusions from this portrayal of the natural history of ALL, assuming for the sake of argument in this review that it is essentially correct: firstly, that childhood ALL has arisen as the major sub-type of paediatric cancer as a paradoxical consequence of our progress—socio-economically and medically speaking, and, secondly, that it may in fact be a preventable disease. Treatment may be successful in 75% or so of cases but it is nasty, toxic and can have serious, morbid sequelae. Some survivors have paid a heavy price. The financial cost of treatment is not trivial either.

The aetiological story for adults with ALL or AML will be different and there are disappointingly few clues to aetiology here, though a minority of cases may be associated with exposure to benzene, other solvents or ionising radiation [68]. For infants with ALL (or AML) and *MLL* gene fusions, it is very likely that different causal mechanisms are at play to those in older children with ALL. Here also there is now a plausible aetiological pathway that is being carefully scrutinised. From the natural history of infant leukaemia, we can conclude that the critical exposures, if any are involved, must occur transplacentally to the fetus during pregnancy. Epidemiological studies have pointed to possible candidate exposures of the pregnant mother (including petrochemical products, solvents, pesticides or alcohol) that are associated with increased risk of leukaemia in the first 2 years of life [28]. Rearrangements of the *MLL* gene are associated in the therapy-associated or secondary acute leukaemias with recent prior exposure to anthracyclines or epidophyllotoxins, drugs that induce DNA cleavage via the formation of cleavable complexes with topo-isomerase II. In situations where cells survive the apoptotic response to such damage, illegitimate

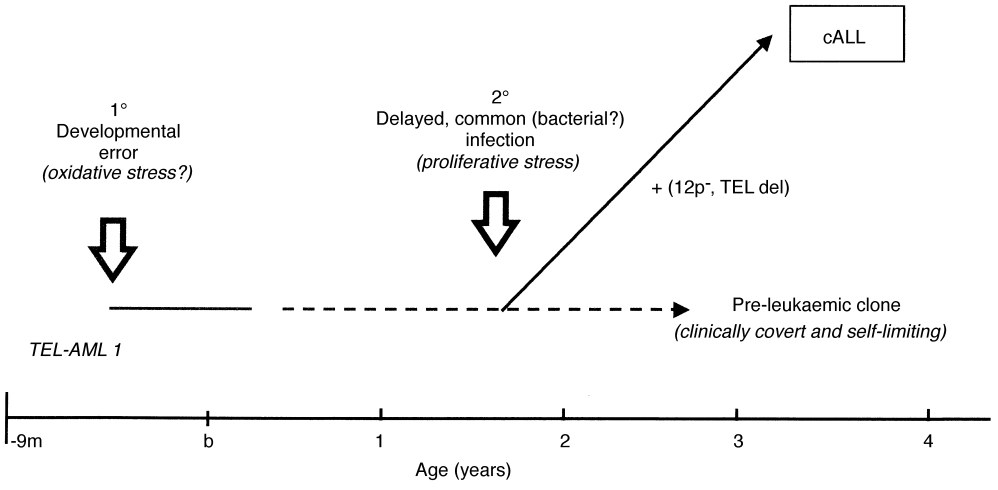


Figure 8. A two-step model for the natural history of cALL.

recombination of genes or chromosomal regions may be facilitated. On this basis it has been suggested that something similar might happen during pregnancy to promote or induce infant leukaemia with *MLL* fusion genes [36, 69]. Pregnant women and their unborn children are sometimes exposed to substances that are topo II inhibitors or can be metabolised to topo II inhibitors, including some antibiotics, laxatives, podophyllin resins, flavonoids in food and drink, herbal medicines and benzene metabolites. These candidate exposures are currently being assessed in case–control epidemiological studies focussed on those leukaemia cases with *MLL* fusion genes. Additionally, since the response to genotoxic insults is modulated by the activity of enzymes that activate or detoxify these chemicals and that are themselves subject to inherited polymorphism, there is a need to consider genetic predisposition or inherited risk in this context. Infant ALL with *MLL* fusion genes have a wretched prognosis [70], few surviving without a matched transplant [1]. Is this cancer preventable?

COLLECTING A SET: CANCER GENES AND CURABILITY

In some cases of paediatric or adult ALL presenting with relatively high white blood cell counts as well as ALL with T lineage phenotypes or ALL in relapse, genetic aberrations in addition to chromosomal translocations are often detectable. These are not leukaemia-specific and include deletions at the p15/p16 (CDK4 inhibitor) loci [71], mutations in *N-RAS* [72] or p53 [73]. Collectively, these are the bad news genotypes, the subset of patients that are unlikely to be cured by conventional chemotherapy. In contrast, these additional molecular abnormalities are almost invariably absent in typical paediatric B cell precursor cALL characterised by low to modest leukaemic burden, excellent initial clinical response and long term outcome. The clinical diversity of acute leukaemia of children and adults is, therefore, intimately tied in with the underlying biological heterogeneity of cellular origin and associated molecular changes [1, 7]. Different individual genetic abnormalities provide clonal advantage for particular cell types that are maximally at risk at different ages. The evolution and natural history of these clonal varieties then differ both in pace and number and character of additional abnormalities accrued by the time that a diagnosis is likely. A full molecular profile of leukaemic blasts at diagnosis would reveal those that are likely either as a population or at the single cell level to be clinically intransigent. Particular fusion genes in paediatric leukaemia, fortunately the minority ones, are bad news—*BCR-ABL* and *MLL* gene fusions, and at least in the former case, this can be rationalised by the pleiotypic activity of the activated ABL kinase [74] including its anti-apoptotic function [75]. Alteration of cell cycle controls is likely to be bad news also as this pathway only endows potent selective advantage if the normally coupled imposition of apoptosis is aborted [76]. Then selection for stress resistance via *p53* loss of function [77] will have deleterious consequences both for immediate therapeutic sensitivity and the risk of accumulating further genetic diversity and drug resistance [78, 79]. Most acute leukaemias will travel through this increasingly selective landscape acquiring clonal fitness if allowed to do so. What seems to be happening in most cases of childhood leukaemia is that the early dissemination of the clone prompts diagnostic symptoms before most of these changes have taken place. The unfortunate cases are those caught late or, less frequently, kick-started by genes that exert more powerful or pleiotypic effects on cell cycle and apoptosis controls.

You can only see what you look for with genetic abnormalities and a great deal of additional molecular shenanigans could still lie below the surface in good prognosis ALL with apparently minimal molecular deviation. If so, this may be revealed by more efficacious and broad genetic screening methods now becoming available—*IGH*, M-FISH, spectral FISH and micro-arrays for example. In the absence of this information, we can entertain the following idea: that it only takes two molecular events to generate the common variant of childhood ALL and in terms of function, this is really the sequential assembly of one biochemical event. The predominant impact may be a block to differentiation which provides for clonal advantage in a bone marrow micro-environment that is already permissive for growth (via appropriate factors) and for migration—in the fetus, neonate and young child. This would explain why primary cALL cells are difficult to clone or maintain *in vitro* and drop dead with any whiff of apoptotic signalling. They are, after all, doing what comes naturally, since normal B lineage progenitors at the stage of *IGH* gene rearrangement are exquisitely sensitive to DNA damage [80]. Then the most effective way to block differentiation would be to corrupt those transcriptional regulators that normally control the process [22]. The same argument applies to many paediatric solid tumours, particularly those of mesenchymal stem cell origin. Knudson has previously suggested that two genetic changes might be sufficient for paediatric cancer in general [81]. If these ideas are correct, then the implications for therapeutic responsiveness are very significant. Fortuitously, ALL and a subset of other cancers originating in embryonic or fetal stem cells, may be able to evolve and disseminate without co-opting the multiple mutant tricks that are necessary for epithelial stem cell progeny similarly to escape and establish a territorial hijack in the adult. These exit cards will include genetic instability, loss of *p53* and CDK inhibitors allowing constitutive cell cycling to be expressed without compensating apotosis [82, 83] or enforced senescence [87]. It is the genetic diversity, relative insensitivity to DNA damage and block to cell suicide that underlies poor clinical response in cancer. Acquiring these credentials may well be an evolutionary inevitability for invasive epithelial cancers. For these cancers also, a block to differentiation may provide less potency for natural clonal selection and initiation. If correct, this would explain the apparent dearth of reciprocal translocations involving genes encoding key transcriptional regulators in carcinomas. In contrast, for ALL and some other paediatric cancers and perhaps also the 'exceptional' (young) adult cancers—testicular cancers, choriocarcinoma and Hodgkin's disease—the developmental biology of the tissues involved provides for a different or more simple pattern of natural selection. Different in terms of types of mutation, number of mutations and time frame from initiation to clinical diagnosis. Somewhat serendipitously, this may have worked very much to our clinical advantage in these 'curable' cancers.

1. Kersey JH. Fifty years of studies of the biology and therapy of childhood leukemia. *Blood* 1997, **90**, 4243–4251.
2. Greaves MF. Differentiation-linked leukaemogenesis in lymphocytes. *Science* 1986, **234**, 697–704.

3. Greaves M. The new biology of leukemia. In Henderson ES, Lister TA, Greaves MF, eds. *Leukemia*. Philadelphia, W B Saunders Company, 1996, 34–45.
4. Heisterkamp N, Stephenson JR, Groffen J, *et al.* Localization of the c-abl oncogene adjacent to a translocation breakpoint in chronic myelocytic leukaemia. *Nature* 1983, **306**, 239–242.
5. Chan LC, Karhi KK, Rayter SI, *et al.* A novel abl protein expressed in Philadelphia chromosome positive acute lymphoblastic leukaemia. *Nature* 1987, **325**, 635–637.
6. Secker-Walker LM. Prognostic and biological importance of chromosome findings in acute lymphoblastic leukaemia. *Cancer Genet Cytogenet* 1990, **49**, 1–13.
7. Greaves MF. Stem cell origins of leukaemia and curability. *Br J Cancer* 1993, **67**, 413–423.
8. Hagemeijer A, Grosveld G. Molecular cytogenetics of leukemia. In Henderson ES, Lister TA, Greaves MF, eds. *Leukemia*. Philadelphia, W B Saunders, 1996, 131–144.
9. Rabbitts TH. Chromosomal translocations in human cancer. *Nature* 1994, **372**, 143–149.
10. Look AT. Oncogenic transcription factors in the human acute leukemias. *Science* 1997, **278**, 1059–1064.
11. Thirman MJ, Gill HJ, Burnett RC, *et al.* Rearrangement of the *MLL* gene in acute lymphoblastic and acute myeloid leukemias with 11q23 chromosomal translocations. *N Engl J Med* 1993, **329**, 909–914.
12. Mrozek K, Bloomfield CD. Major cytogenetic findings in non-Hodgkin's lymphoma. In Canellos GP, Lister TA, Sklar JL, eds. *The Lymphomas*. Philadelphia, W B Saunders, 1998, 107–128.
13. Greaves MF. Workshop Report. Infant leukaemia biology, aetiology and treatment. *Leukemia* 1996, **10**, 372–377.
14. Raimondi SC. Current status of cytogenetic research in childhood acute lymphoblastic leukemia. *Blood* 1993, **81**, 2237–2251.
15. Romana SP, Poirel H, Leconiat M, *et al.* High frequency of t(12;21) in childhood B-lineage acute lymphoblastic leukemia. *Blood* 1995, **86**, 4263–4269.
16. Secker-Walker LM, Craig JM, Hawkins JM, Hoffbrand AV. Philadelphia positive acute lymphoblastic leukemia in adults: age distribution, BCR breakpoint and prognostic significance. *Leukemia* 1991, **5**, 196–199.
17. Rowley JD. Identification of the constant chromosome regions involved in human hematologic malignant disease. *Science* 1982, **216**, 749–750.
18. Greaves M. A natural history for pediatric acute leukemia. *Blood* 1993, **82**, 1043–1051.
19. Barr FG. Translocations, cancer and the puzzle of specificity. *Nature Genet* 1998, **19**, 121–124.
20. Griffiths SD, Healy LE, Ford AM, *et al.* Clonal characteristics of acute lymphoblastic cells derived from *BCR/ABL* p190 transgenic mice. *Oncogene* 1992, **7**, 1391–1399.
21. Corral J, Lavenir I, Impey H, *et al.* An Mll-AF9 fusion gene made by homologous recombination causes acute leukemia in chimeric mice: a method to create fusion oncogenes. *Cell* 1996, **85**, 853–861.
22. Enver T, Greaves M. Loops, lineage, and leukemia. *Cell* 1998, **94**, 9–12.
23. Taylor GM, Birch JM. The Hereditary Basis of Human Leukemia. In Henderson ES, Lister TA, Greaves MF, eds. *Leukemia*. Phildelphia, W B Saunders Company, 1996, 210–245.
24. Robison LL, Nesbit ME, Sather HN, *et al.* Down syndrome and acute leukemia in children: a 10-year retrospective survey from Childrens Cancer Study Group. *J Pediatr* 1984, **105**, 235–242.
25. Heim S, Bekassy AN, Garwicz S, *et al.* New structural chromosomal rearrangements in congenital leukemia. *Leukemia* 1987, **1**, 16–23.
26. Ridge SA, Cabrera ME, Ford AM, *et al.* Rapid intraclonal switch of lineage dominance in congenital leukaemia with a *MLL* gene rearrangement. *Leukemia* 1995, **9**, 2023–2026.
27. Hunger SP, McGavran L, Meltesen L, Parker NB, Kassenbrock CK, Bitter MA. Oncogenesis *in utero* fetal death due to acute myelogenous leukaemia with an *MLL* translocation. *Br J Haematol* 1998, **103**, 539–542.
28. Ross JA, Davies SM, Potter JD, Robison LL. Epidemiology of childhood leukemia, with a focus on infants. *Epidemiol Rev* 1994, **16**, 243–272.
29. Smith MA, Chen T, Simon R. Age-specific incidence of acute lymphoblastic leukemia in U.S. children: *in utero* initiation model. *J Natl Cancer Inst* 1997, **89**, 1542–1544.
30. Strong SJ, Corney G. *The Placenta in Twin Pregnancy*. Pergamon Press, Oxford, 1967.
31. Kumar D, Gemayel NS, Deapen D, *et al.* North-American twins with IDDM. Genetic, etiological and clinical significance of disease concordance according to age, zygosity, and the interval after diagnosis in first twin. *Diabetes* 1993, **42**, 1351–1363.
32. Sadovnick AD, Armstrong H, Rice GPA, *et al.* A population-based study of multiple sclerosis in twins: update. *Ann Neurol* 1993, **33**, 281–285.
33. MacMahon B, Levy MA. Prenatal origin of childhood leukemia. *N Engl J Med* 1964, **270**, 1082–1085.
34. Danis RP, Keith LG. Some observations concerning leukemia in twins. *Acta Genet Med Gemellol* 1982, **31**, 173–177.
35. Clarkson B, Boyse EA. Possible explanation of the high concordance for acute leukaemia in monozygous twins. *Lancet* 1971, **i**, 699–701.
36. Ford AM, Ridge SA, Cabrera ME, *et al. In utero* origin of rearrangements in the trithorax-related oncogene in infant leukaemias. *Nature* 1993, **363**, 358–360.
37. Gill Super HJ, Rothberg PG, Kobayashi H, Freeman AI, Diaz MO, Rowley JD. Clonal, nonconsitutional rearrangements of the *MLL* gene in infant twins with acute lymphoblastic leukemia: *in utero* chromosome rearrangement of 11q23. *Blood* 1994, **83**, 641–644.
38. Megonigal MD, Rappaport EF, Jones DH, *et al.* t(11;22)(q23; q11.2) in acute myeloid leukemia of infant twins fuses *MLL* with hCDC*rel*, a cell division cycle gene in the genomic region of deletion in DiGeorge and velocardiofacial syndromes. *Proc Natl Acad Sci USA* 1998, **95**, 6413–6418.
39. Ford AM, Pombo-de-Oliveira MS, McCarthy KP, *et al.* Monoclonal origin of concordant T-cell malignancy in identical twins. *Blood* 1997, **89**, 281–285.
40. Ford AM, Bennett CA, Price CM, Bruin MCA, Van Wering ER, Greaves M. Fetal origins of the *TEL-AML1* fusion gene in identical twins with leukemia. *Proc Natl Acad Sci USA* 1998, **95**, 4584–4588.
41. Wiemels JL, Ford AM, Van Wering ER, Postma A, Greaves M. Protracted and variable latency of acute lymphoblastic leukaemia after *TEL-AML1* gene fusion *in utero*. *Blood* 1999, in press.
42. Wiemels JL, Ford AM, Bennett CA, *et al.* Direct evidence for a pre-natal origin of *TEL-AML1* gene fusions in paediatric leukaemia. *Blood* 1998, **92**(Suppl. 1), 68.
43. Jinks DC, Minter M, Tarver DA, Vanderford M, Hejtmancik JF, McCabe ERB. Molecular genetic diagnosis of sickle cell disease using dried blood specimens on blotters used for newborn screening. *Hum Genet* 1989, **81**, 363–366.
44. Raskin S, Phillips JA, Kaplan G, McClure M, Vnencak-Jones C. Cystic fibrosis genotyping by direct PCR analysis of Guthrie blood spots. *PCR Methods and Applications* 1992, **2**, 154–156.
45. Comeau AM, Hsu H-W, Schwerzler M, *et al.* Identifying human immunodeficiency virus infection at birth application of polymerase chain reaction to Guthrie cards. *J Pediatr* 1993, **123**, 252–258.
46. Noda S, Eizuru Y, Minamishima Y, Ikenoue T, Mori N. Detection of human T-cell lymphotropic virus type 1 infection by the polymerase chain reaction using dried blood specimens on filter papers. *J Virol Methods* 1993, **43**, 111–122.
47. Gale KB, Ford AM, Repp R, *et al.* Backtracking leukemia to birth: identification of clonotypic gene fusion sequences in neonatal blood spots. *Proc Natl Acad Sci USA* 1997, **94**, 13950–13954.
48. Smith MA, McCaffrey RP, Karp JE. The secondary leukemias: Challenges and research directions. *J Natl Cancer Inst* 1996, **88**, 407–418.
49. Romana SP, Mauchauffe M, Le Coniat M, *et al.* The t(12;21) of acute lymphoblastic leukemia results in a *tel-AML1* gene fusion. *Blood* 1995, **85**, 3662–3670.
50. Raynaud S, Cave H, Baens M, *et al.* The 12;21 translocation involving *TEL* and deletion of the other *TEL* allele: two frequently associated alterations found in childhood acute lymphoblastic leukemia. *Blood* 1996, **87**, 2891–2899.
51. Wang LC, Swat W, Fujiwara Y, *et al.* The *TEL/ETV6* gene is required specifically for hematopoiesis in the bone marrow. *Genes Dev* 1998, **12**, 2392–2402.
52. Okuda T, van Deursen J, Hiebert SW, Grosveld G, Downing JR. AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. *Cell* 1996, **84**, 321–330.

53. Hiebert SW, Sun W, Davis JN, *et al*. The t(12;21) translocation converts AML-1B from an activator to a repressor of transcription. *Mol Cell Biol* 1996, **16**, 1349–1355.
54. McLean TW, Ringold S, Neuberg D, *et al*. TEL/AML-1 dimerizes and is associated with a favorable outcome in childhood acute lymphoblastic leukemia. *Blood* 1996, **88**, 4252–4258.
55. Greaves MF, Alexander FE. An infectious etiology for common acute lymphoblastic leukemia in childhood? *Leukemia* 1993, **7**, 349–360.
56. Kinlen LJ. Epidemiological evidence for an infective basis in childhood leukaemia. *Br J Cancer* 1995, **71**, 1–5.
57. Greaves MF. Aetiology of acute leukaemia. *Lancet* 1997, **349**, 344–349.
58. MacKenzie J, Perry J, Ford AM, Jarrett RF, Greaves M. JC and BK virus sequences are not detectable in leukemic samples from children with common acute lymphoblastic leukemia. *Br J Cancer* 1999, in press.
59. Tsiagbe VK, Yoshimoto T, Asakawa J, Cho SY, Meruelo D, Thorbecke GJ. Linkage of superantigen-like stimulation of syngeneic T cells in a mouse model of follicular center B cell lymphoma to transcription of endogenous mammary tumor virus. *EMBO J* 1993, **12**, 2313–2320.
60. Wotherspoon AC, Doglioni C, Diss TC, *et al*. Regression of primary low-grade B-cell gastric lymphoma of mucosa-associated lymphoid tissue type after eradication of *Helicobacter pylori*. *Lancet* 1993, **342**, 575–577.
61. Poynton F J, Thursfield H, Paterson D. The severe blood diseases of childhood: a series of observations from the Hospital for Sick Children, Great Ormond Street. *Br J Child Dis* 1922, **XIX**, 128–144.
62. van Steensel-Moll HA, Valkenburg HA, van Zanen GE. Childhood leukemia and infectious diseases in the first year of life a register-based case/control study. *Am J Epidemiol* 1986, **124**, 590–594.
63. Petridou E, Kassimos D, Kalmanti M, *et al*. Age of exposure to infections and risk of childhood leukaemia. *Br Med J* 1993, **307**, 774.
64. Shu XO, Wen W, Buckley J, Brondum J, Linet M, Robison L. Lifestyle and risk of acute lymphocytic leukemia in children. *Am J Epidemiol* 1998, **147**(Suppl.), 37.
65. Tycko B, Sklar J. Chromosomal translocations in lymphoid neoplasia: a reappraisal of the recombinase model. *Cancer Cells* 1990, **2**, 1–8.
66. Wiemels JL, Greaves MF. Genomic characterisation of the *TEL-AML1* fusion in childhood acute lymphoblastic leukaemia. *Cancer Res* 1999, in press.
67. Zucman-Rossi J, Legoix P, Victor J-M, Lopez B, Thomas G. Chromosome translocation based on illegitimate recombination in human tumors. *Proc Natl Acad Sci USA* 1998, **95**, 11786–11791.
68. Groves FD, Linet MS, Devesa SS. Epidemiology of leukemia: overview and patterns of occurrence. In Henderson ES, Lister TA, Greaves MF, eds. *Leukemia*. Philadelphia, W B Saunders, 1996, 145–159.
69. Ross JA, Potter JD, Robison LL. Infant leukemia, topoisomerase II inhibitors, and the *MLL* gene. *J Natl Cancer Inst* 1994, **86**, 1678–1680.
70. Heerema NA, Arthur DC, Sather H, *et al*. Cytogenetic features of infants less than 12 months of age at diagnosis of acute lymphoblastic leukemia: impact of the 11q23 breakpoint on outcome: a report of the Children's Cancer Group. *Blood* 1994, **83**, 2274–2284.
71. Takeuchi S, Bartram CR, Seriu T, *et al*. Analysis of a family of cyclin-dependent kinase inhibitors p15/MTS2/INK4B, p16/MTS1/INK4A, a p18 genes in acute lymphoblastic leukemia of childhood. *Blood* 1995, **86**, 755–760.
72. Rodenhuis S, Bos JL, Slater RM, Behrendt H, van't Veer M, Smets LA. Absence of oncogene amplifications and occasional activation of N-*ras* in lymphoblastic leukemia of childhood. *Blood* 1986, **67**, 1698–1704.
73. Wada M, Bartram CR, Nakamura H, *et al*. Analysis of p53 mutations in a large series of lymphoid hematologic malignancies of childhood. *Blood* 1993, **82**, 3163–3169.
74. Goga A, McLaughlin J, Afar DEH, Saffran DC, Witte ON. Alternative signals to RAS for hematopoietic transformation by the *BCR-ABL* oncogene. *Cell* 1995, **82**, 981–988.
75. Kabarowski JHS, Allen PB, Wiedemann LM. A temperature sensitive p210 *BCR-ABL* mutant defines the primary consequences of *BCR-ABL* tyrosine kinase expression in growth factor dependent cells. *EMBO J* 1994, **13**, 5887–5895.
76. Harrington E, Fanidi A, Evan G. Oncogenes and cell death. *Curr Opin Genet Dev* 1994, **4**, 120–129.
77. Kinzler KW, Vogelstein B. Life (and death) in a malignant tumor. *Nature* 1996, **379**, 19–20.
78. Lowe SW, Bodis S, McClatchey A, *et al*. *p53* status and the efficacy of cancer therapy *in vivo*. *Science* 1994, **266**, 807–810.
79. Griffiths SD, Clarke AR, Healy LE, *et al*. Absence of p53 permits propagation of mutant cells following genotoxic damage. *Oncogene* 1997, **14**, 523–531.
80. Griffiths SD, Goodhead DT, Marsden SJ, *et al*. Interleukin 7-dependent B lymphocyte precursor cells are ultrasensitive to apoptosis. *J Exp Med* 1994, **179**, 1789–1797.
81. Knudson AG. Stem cell regulation, tissue ontogeny, and oncogenic events. *Sem Cancer Biol* 1992, **3**, 99–106.
82. Fearon E, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell* 1990, **61**, 759–767.
83. Hunter T, Pines J. Cyclins and cancer II: cyclin D and CDK inhibitors come of age. *Cell* 1994, **79**, 573–582.
84. Mack TM, Cozen W, Shibata DK, *et al*. Concordance for Hodgkin's disease in identical twins suggesting genetic susceptibility to the young-adult form of the disease. *N Engl J Med* 1995, **332**, 413–418.
85. Zuelzer WW, Cox DE. Genetics aspects of leukemia. *Sem Hematol* 1969, **228**, 228–249.
86. Buckley JD, Buckley CM, Breslow NE, Draper GJ, Roberson PK, Mack TM. Concordance for childhood cancer in twins. *Med Pediatr Oncol* 1996, **26**, 223–229.
87. Serrano M, Lin AW, McCurrach ME, Beach D, Lowe SW. Oncogenic *ras* provokes premature cell senescence associated with accumulation of p53 and $p16^{INK4a}$. *Cell* 1997, **88**, 593–602.

Acknowledgements—The author holds a Leukaemia Research Fund specialist programme. Additional support has been provided by the Kay Kendall Leukaemia Fund. I am grateful to those colleagues who have contributed to data referred to in the text, to clinicians who provided samples from patients, to Dr Leanne Wiedemann for discussion and advice and to Barbara Deverson for help in producing the manuscript.