TITLE: Timing, Sequence and Sub-clonal Architecture of Genetic Events in Acute Lymphoblastic Leukemia

(*The original project proposal has been significantly amended since the outlined study was published by a competitor research group just at the beginning of the fellowship period.)*

INTRODUCTION

Acute Lymphoblastic Leukemia (ALL) is the most common malignancy in childhood. Concordance of ALL in monozygotic twins derives from intraplacental vascular anastomoses and blood cell chimeras and consequent sharing of pre-leukemic clones initiated in one of the twins only. Thus, leukemias in these twins have a common *in utero* origin which is followed by a 1 to 14-year-long independent leukemogenesis, from pre-leukemic stage to overt leukemia, in the siblings after birth. Consequently, twins with concordant or discordant ALL provide a unique angle to study the timing, sequence and subclonal dynamics of leukemia-associated genetic changes.

AIMS/OBJECTIVES

Our project aimed

1) To reveal the developmental stage of the cell from which the malignant clonal expansion originates in *ETV6/RUNX1* gene fusion positive and hyperdiploid B-cell precursor ALLs.
2) To compare the evolutionary trajectories and subclonal genetic architectures of tumour cell populations in monozygotic twin children with concordant B-cell precursor ALL.
3) To reveal if the putative pre-leukemic cells in the healthy co-twin of patients with *ETV6/RUNX1* gene fusion positive ALL are functional, i.e. they have the same genomic fusion gene sequence and express the abnormal fusion transcript as the twin with leukemia.

MAIN RESULTS

Immunoglobulin (*IG*) and T-cell receptor (*TCR*) gene rearrangements, which are hallmarks of normal lymphoid cell development, have comprehensively been screened in diagnostic DNA samples of 5 twin pairs with *ETV6/RUNX1* gene fusion-positive and 2 twin pairs with high-hyperdiploid (chromosome number 51-67 per cell) ALL. These two largest subtypes of the disease comprise half of all cases diagnosed with B-cell precursor ALL. All twin pairs were analyzed using a conventional, polymerase chain reaction (PCR)-based method, and 3 twin pairs harboring *ETV6/RUNX1* gene fusion have also been screened using a novel next-generation sequencing approach that allowed the unbiased analysis of all *IG* and *TCR* gene sequences, simultaneously. Altogether, over 100 clonal *IG/TCR* rearrangements were identified in the 7 twin pairs. Those rearrangements which were found in one patient only were cross-tested in his/her co-twin using highly sensitive allele-specific PCR assays. We observed 1-3 identical genomic rearrangements, also including a number of randomly incorporated nucleotides, shared by the siblings in each twin pair. These markers present in both twins at diagnosis must have belonged to the pre-leukemic cell population formed *in utero*. The vast majority of *IG/TCR* rearrangements proved to be patient-specific thus occurred secondary and probably postnatally. Evaluation of *IG/TCR* profiles in the context of monozygotic twins suggests an early pro- or pre-B cell origin of pre-leukemic clonal expansion in both *ETV6/RUNX1*-positive and high-hyperdiploid ALL.

By combining results of *IG/TCR* analyses with genome-wide copy number profiling, fluorescence *in situ* hybridization, targeted sequencing and exome sequencing data, we have revealed evolutionary trajectories of high-hyperdiploid ALL, which suggest a B-cell precursor *in utero* origin followed by subclonal and mostly postnatal accumulation of driver mutational changes necessary for clinical manifestation in this subtype of ALL. *Bateman*, Alpar* et al. Leukemia 10.1038/leu.2014.177*.

In one pair of twins with *ETV6/RUNX1*+ ALL, we performed single-cell genotyping using a high-throughput microfluidic quantitative PCR platform allowing the combined investigation of genetic copy number abnormalities, single nucleotide variants, gene fusions and *IG/TCR*

rearrangements in single cells unbiasedly selected from the whole leukemia. *ETV6-RUNX1* gene fusion, *ETV6* gene loss, *RUNX1* gene gain, 9 coding-region single nucleotide variants and 8 *IG/TCR* rearrangements were included in the analysis. Five subclones with 9-10 genetic markers and 8 subclones with 7-15 markers have been identified in twins A and B, respectively. Branching, divergent subclonal architectures with different complexity were observed in the siblings by reconstruction of the most likely phylogenetic relationships between the detected subclones. Copy number aberrations occurred multiple times on independent branches of the phylogenetic tree which suggests convergent evolution. Interestingly, the most evolved subclone harboring the highest number of aberrations did not seem to be the most aggressive one, at least based on this snapshot analysis. Combined analysis of all subclones detected in the twins demonstrated that immunogenotype diversification can start before birth and multiple *in utero* formed subclones, originating from the same pre-leukemic clone but having independent immunogenotype, can survive for years and contribute to the development of clinically overt leukemia.

Investigating two twin pairs *dis*cordant for B-cell precursor ALL, we identified putative pre-leukemic cells harboring the *ETV6/RUNX1*-fusion in a proportion of less than 1% of the B-cell compartment of healthy twins using fluorescence in situ hybridization. Genome-level scrutiny revealed identical *ETV6/RUNX1* fusion breakpoints in bulk DNA isolated from B-cells of twin siblings, providing evidence of the presence of prenatally occurred *ETV6/RUNX1*-positive cells. *ETV6/RUNX1* transcripts have been detected in bulk RNA isolated from sorted B cells, thought to have self-renewing ability in ALL. Although, the single-cell analysis of *ETV6/RUNX1* transcripts proved to be challenging due to technical limitations, these data indirectly suggest that pre-leukemic cells in the healthy co-twin of patients with *ETV6/RUNX1* positive B-cell precursor ALL are functional and express the abnormal fusion transcript.

CONCLUSIONS

Our data suggest that (*i*) the pre-leukemic initiating function of *ETV6-RUNX1* fusion is associated with clonal expansion early in the fetal B cell lineage and similar results were obtained in high-hyperdiploid ALL; (*ii*) leukemias originating in a common *in utero* environment can have divergent subclonal architectures with different complexity; (*iii*) multiple *in utero* formed pre-leukemic subclones evolve and can gain selective advantage in parallel postnatally and (*iv*) *ETV6/RUNX1*-positive pre-leukemic cell populations with putative self-renewing ability are functional and express the fusion transcript.

INNOVATION

In collaboration with our industrial partner, the MRC-Holland (Amsterdam), we worked on the development of a multiple ligation-dependent probe amplification (MLPA) kit which allows the rapid and affordable analysis of genetic copy number abnormalities associated with T-cell ALL. In the frame of this separate sub-project, a new MLPA mix containing 54 probes for 13 different chromosomal regions, suggested to be of diagnostic and/or prognostic importance in T-ALL, has been introduced. To test the efficiency of the method, 52 diagnostic bone marrow samples have been screened. MLPA detected all CNAs previously identified by single nucleotide polymorphism (SNP)-array and also revealed small losses/gains which were missed by SNP-array.

SOCIO-ECONOMIC IMPACT

After accidents, ALL is the second leading cause of death among children between ages 1 and 14 in the Western World. Although, therapy response and survival rate have substantially improved recently, a significant number of patients still suffers from late relapses and serious side effects. Our study has contributed to the in-depth understanding of the genetic basis of this malignancy and the results will probably influence future clinical management strategies from which many children with cancer can benefit.