LUND UNIVERSITY

Genetic and epigenetic characterization of pediatric high hyperdiploid acute lymphoblastic leukemia

Davidsson, Josef

2009

GENETIC AND EPIGENETIC CHARACTERIZATION OF PEDIATRIC HIGH HYPERDIPLOID ACUTE LYMPHOBLASTIC LEUKEMIA

Akademisk Avhandling

av

JOSEF DAVIDSSON

Blekingska Nationen

som med vederbörligt tillstånd av Medicinska Fakulteten vid Lunds universitet
för avläggande av doktorsexamen i medicinsk vetenskap kommer att
offentligen försvaras i föreläsningssal F1, centralblocket, Universitetssjukhuset
i Lund, fredagen den 27 november 2009 kl 10.00

Fakultetsopponent:

Dr Lyndal Kearney
Section of Haemato-Oncology
The Institute of Cancer Research, University of London
Sutton, UK

Organization	Document name
LUND UNIVERSITY Department of Clinical Genetics Lund University Hospital SE-221 85 Lund	DOCTORAL DISSERTATION
	Date of issue 27th of November 2009
	Sponsoring organization The Swedish Childhood Cancer Foundation The Swedish Cancer Society The Swedish Research Council
Author(s) Josef Davidsson	

Title and subtitle

Genetic and epigenetic characterization of pediatric high hyperdiploid acute lymphoblastic leukemia

Abstract

The aim of this thesis was to analyze the genetic and epigenetic characteristics of pediatric high hyperdiploid acute lymphoblastic leukemia (HeH ALL), the most common type of childhood malignancy. The three original articles presented in this thesis have addressed three major questions regarding HeH ALL: what are the genetic characteristics of the most common structural abnormality – dup(1q)? Are aberrant methylation patterns involved in leukemogenesis? How do diagnostic and relapse samples relate to each other genetically? In article I we found that the proximal breakpoints of dup(1q)-positive Burkitt lymphomas (BL) and pediatric HeH ALLs cluster close to the centromere, suggesting this region to be breakprone, and that five genes on 1q were significantly overexpressed and thus potentially pathogenetically important. In addition, the satellite II domain was not hypomethylated in either disorder, ruling out that decondensation of pericentric heterochromatin, due to hypomethylation of satellite II sequences, is the mechanism behind dup(1q) in BL and pediatric HeH ALL. In article II unsupervised cluster and principal component analyses of BAC array chromosome-wide methylome profiles accurately subgrouped the t(12;21) and HeH ALL subtypes, demonstrating that they are characterized by subtype specific methylomes. Analysis of promoter-specific CpG islands demonstrated that several B-cell- and neoplasia-associated genes were hypermethylated and underexpressed. Methylation hotspots were associated with chromosome bands predicted to harbor imprinted genes, suggesting that regions rich in imprinted genes could be associated with aberrant methylation in t(12;21)-positive and HeH ALLs. Moreover, we demonstrated that the tri-/tetrasomic chromosomes in HeH ALLs were less methylated than their disomic counterparts. Decreased methylation of gained chromosomes is a previously unknown phenomenon that may have ramifications not only for the pathogenesis of HeH ALL but also for other disorders with acquired or constitutional numerical chromosome anomalies. In article III we concluded that structural changes were significantly more common in relapse HeH samples than in diagnostic ones and that no single, recurrent aberration could be directly linked to relapse. In addition, the triple trisomies +4, +10, and +17, associated with low-risk ALL, were detected in two-thirds of cases that subsequently relapsed, thus questioning the favorable impact of these trisomies. The genetic differences between diagnostic and relapse samples suggest the presence of an ancestral HeH clone in the vast majority of patients, with strucutral changes and mutations being secondary to the HeH pattern.

Key words: high hyperdiploid ALL, children, chromosome aberration, methylation profiling, relapse, array CGH, SNP array, clonal relationship, dup(1q)

Classification system and/or index termes (if any):

Supplementary bibliographical information:		Language English
ISSN and key title: 1652-8220		ISBN 978-91-86443-01-6
Recipient's notes	Number of pages 186	Price
	Security classification	

Genetic and Epigenetic Characterization of Pediatric High Hyperdiploid Acute Lymphoblastic Leukemia

Josef Davidsson

Department of clinical genetics

Faculty of medicine

Lund university

2009

This thesis is dedicated to my mother and father,
for always believing in me.

What is it that we human beings ultimately depend on? We depend on our words. We are suspended in language. Our task is to communicate experience and ideas to others.

Niels Bohr

ISSN 1652-8220

ISBN 978-91-86443-01-6

Printed in Sweden by Media Tryck AB, Lund 2009

TABLE OF CONTENTS

ABBREVIATIONS

ALL	acute lymphoblastic leukemia
AML	acute myeloid leukemia
BAC	bacterial artificial chromosome
BASE	bioarray software environment
BL	Burkitt lymphoma
CGH	comparative genomic hybridization
CML	chronic myeloid leukemia
COG	Children's Oncology Group
FISH	fluorescent *in situ* hybridization
GWA	genome-wide association
HeH	high hyperdiploid
MeDIP	methylated DNA immunoprecipitation
NOPHO	Nordic Society of Paediatric Haematology and Oncology
PCR	polymerase chain reaction
Ph	Philadelphia chromosome
RT	reverse transcriptase
SNP	single nucleotide polymorphism
UPD	uniparental isodisomy
WBC	white blood cell

ORIGINAL ARTICLES

This thesis is based on the following articles, which will be referred to in the text by their Roman numerals as listed below.

I. Davidsson J, Andersson A, Paulsson K, Heidenblad M, Isaksson M, Borg Å, Heldrup J, Behrendtz M, Panagopoulos I, Fioretos T and Johansson B (2007). Tiling resolution array comparative genomic hybridization, expression and methylation analyses of dup(1q) in Burkitt lymphomas and pediatric high hyperdiploid acute lymphoblastic leukemias reveal clustered near-centromeric breakpoints and overexpression of genes in 1q22-32.3. *Hum Mol Genet* **16:** 2215-2225.

II. Davidsson J, Lilljebjörn H, Andersson A, Veerla S, Heldrup J, Behrendtz M, Fioretos T and Johansson B (2009). The DNA methylome of pediatric acute lymphoblastic leukemia. *Hum Mol Genet* **18:** 4054-4065.

III. Davidsson J, Paulsson K, Lindgren D, Lilljebjörn H, Chaplin T, Forestier E, Andersen MK, Nordgren A, Rosenquist R, Fioretos T, Young BD and Johansson B (2009). Relapsed childhood high hyperdiploid acute lymphoblastic leukemia: presence of preleukemic ancestral clones and the secondary nature of microdeletions and RTK-RAS mutations. *Submitted.*

PREFACE

The treatment of childhood acute lymphoblastic leukemia (ALL) is one of the true success stories of modern medicine. Before 1950, childhood leukemia was uniformly fatal, usually within a period of 3 months. Death from hemorrhages and severe infections was routine and the only treatment available, blood transfusions, offered little help. However, during the 1950's, treatment radically changed with the introduction of drugs like methotrexate, 6-mercaptopurine and cortisone; chemotherapy had been born. Since 1975, the refinements of existing protocols have resulted in a survival rate of nearly 80%. Although this is remarkable, many problems linger. Except for Imatinib, used to treat Philadelphia chromosome (Ph) positive ALL, no new major therapeutic breakthrough has been made during the last 30 years and therapy still remains complex, expensive and very toxic to the patient. In addition, the disease etiology is unknown and there is thus no way of preventing ALL.

The topic of this thesis is the most common form of pediatric ALL, namely high hyperdiploid (HeH) ALL, a malignancy with a favorable outcome, but unknown pathogenesis. The thesis is divided in four sections; the first introducing hematologic malignancies. In the second, section a review of pediatric HeH ALL is given. The third comprises the aims of the thesis, the material and methods used, results and a discussion, as well as the conclusions drawn. The final section includes the articles upon which the thesis is based.

Lund, October 2009

INTRODUCTION

Hematologic malignancies

Hematologic malignancies comprise all the clinically, morphologically and immunophenotypically heterogeneous neoplastic disorders of the bone marrow, involving both the myeloid and lymphoid lineages of the hematopoietic system. Acute proliferative disorders, known as either acute myeloid leukemia (AML) or ALL depending on their lineage, are characterized by an accumulation of malignant immature white blood cells (blasts) in the bone marrow and, in most cases, also the peripheral blood. This causes dysfunction of the normal hematopoiesis resulting in anemia, leukocytopenia and thrombocytopenia and leading to infection and excessive bleeding in the patient (Jandl, 1996).

Leukemia is a malignant disorder with a clonal origin from one single hematopoietic progenitor cell that has acquired somatic genetic changes causing neoplastic transformation. At present, the diagnosis and prognosis of the acute leukemias are based on the white blood cell (WBC) count, age of the patient, immunophenotype and morphology of the blasts and the presence of specific genetic abnormalities in the neoplastic cells (Pui and Evans, 2006). According to the Association of the Nordic Cancer Registries approximately 1000 cases of acute leukemia occur in the Nordic countries each year (http://www.ancr.nu).

AML predominates in adults, whereas ALL primarily is a pediatric disease (Jandl, 1996). With a characteristic age peak around 2-5 years (Pui *et al.*, 2008) and an incidence of 5 cases per 100 000 and year in the Nordic countries (Hjalgrim *et al.*, 2003a), ALL constitutes the most common childhood malignancy. Therefore, both when investigating pathogenesis as a scientist and when managing the leukemia in the clinic as a practitioner, ALL primarily pose a pediatric challenge. Moreover, this special setting as a childhood disease is important to keep in mind also when investigating the genetic and epigenetic origin of ALL, which is the topic of the present thesis.

Genetic aberrations in ALL

Acquired genetic aberrations have, during the last decades, been catalogued and associated with so many tumor types that they now count by the thousands (Mitelman *et al.*, 2009). In some sense, it could be argued that "the riddle of cancer" to a large extent is solved. The evidence that neoplasia essentially is a genetic disease caused by a dysregulation of cellular systems due to somatic mutations in one or many founder cells is now overwhelming (Hanahan and Weinberg, 2000). This knowledge has had a major impact on the clinical management of patients with ALL and has played a significant role in helping us understand the biology underlying leukemogenesis.

Historically, and also at present, the most characteristic cytogenetic feature of hematologic malignancies has been balanced chromosome translocations (**Figure 1A**) (Rowley, 2008). Two main functional consequences are associated with chromosome translocations. First, a balanced rearrangement can place a gene's coding sequences under the transcriptional control of a regulatory element in close vicinity of the other

breakpoint, causing overexpression of the sequence from the first mentioned gene. Second, a translocation may result in a fusion of two different genes, generating chimeric transcripts with novel functions in the cell (Rabbitts and Stocks, 2003). Other cytogenetic aberrations commonly found in hematologic malignancies are inversions, deletions and duplications (**Figure 1B-D**). Inversions are associated with the same functional outcome as translocations, whereas duplications and deletions traditionally are regarded as imbalances causing gene copy-number gain or loss on a whole chromosome- or individual segment level. Moreover, in many different hematologic malignancies somatic point mutations have been described for a wide variety of genes, causing both gene up-regulation by activating mutations or haploinsufficiency by loss-of-function mutations (Bamford *et al.*, 2004).

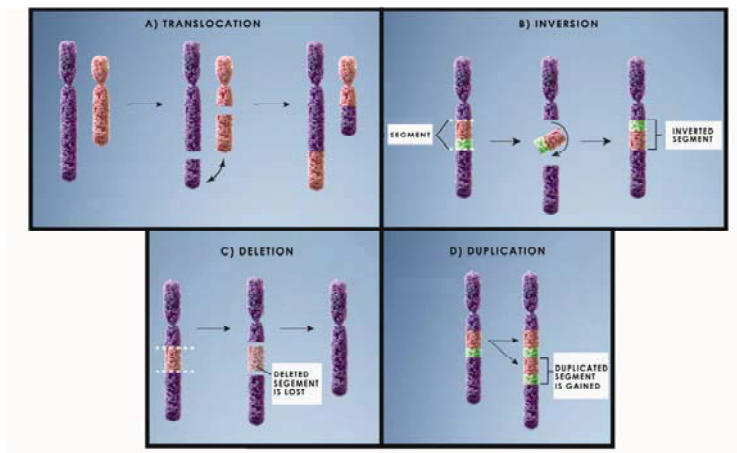


Figure 1. Common chromosome abnormalities in neoplastic disorders. A) A translocation is a reciprocal exchange of genetic material between two chromosomes. B) An inversion is the reversal of position of a chromosomal segment. C) A deletion is when a part of the chromosome is lost, resulting in copy-number loss. D) A duplication is when a part of the chromosome is repeated, resulting in copy-number gain.

Cytogenetic groups may also be classified according to ploidy, based on the number of chromosomes the tumor cell population harbors (Shaffer *et al.*, 2009). In ALL, they comprise diploidy (46 chromosomes), HeH (51-67 chromosomes), near-haploidy (25-29 chromosomes), low hypodiploidy (31-39 chromosomes), near-triploidy (66-79 chromosomes) and near-tetraploidy (84-100 chromosomes) (Harrison and Johansson, 2009). HeH is by far the most common aberrant group in ALL; in fact, it is the overall most common cytogenetic group, also when including subgroups defined by specific structural aberrations. The other aneuploidy groups are rather rare entities in ALL when making the same comparison (Johansson *et al.*, 2004). The functional outcome of these massive chromosomal imbalances has long been debated, something that I will return to in this thesis.

The chromosome morphology is known to be notoriously poor in ALL. In spite of this, structural chromosome abnormalities are detected in approximately 50% of all cases (Harrison and Foroni, 2002), a number that is likely to rise with the entry of novel high resolution techniques into research and clinics. The discovery of recurring chromosomal abnormalities associated with patient outcome has both revealed key genes involved in leukemogenesis as well as enabled cytogenetic subclassification of ALL. Today, such cytogenetic classification contributes to risk stratification of the patients into different treatment regimens (Pui and Evans, 2006; Pui *et al.*, 2008). The most common cytogenetic subgroups in ALL are shortly reviewed below (HeH will be discussed separately):

The t(12;21)(p13;q22) translocation generating the *ETV6/RUNX1* fusion occurs in about 25% of all pediatric ALLs and the resulting protein seems to lead to hampered early lymphocyte development (Pui *et al.*, 2008).

This aberration is not visible by conventional cytogenetics and hence have to be detected by fluorescent *in situ* hybridization (FISH) or reverse transcriptase (RT) polymerase chain reaction (PCR) (Romana *et al.*, 1994; Golub *et al.*, 1995). The outlook for patients harboring this translocation has been considered favorable; however, some studies indicate that this abnormality is associated with late relapses (Harbott *et al.*, 1997; Forestier *et al.*, 2008a). Secondary aberrations, such as deletions of 6q21-27, 8p11-23, 9p13-24, 11q23-25, 12p11-13, 13q14-34 and the whole X chromosome and gains of chromosomes 10, 16 and 21 are common in t(12;21)-positive ALL but do not seem to affect prognosis (Forestier *et al.*, 2007).

t(1;19)(q23;p13) [*TCF3/PBX1*] occurs in a balanced form as well as an unbalanced der(19)t(1;19). This translocation is found in approximately 5% of all B cell precursor ALLs and the resulting protein product leads to cell differentiation arrest (Hunger, 1996). For pediatric patients this aberration is risk stratifying in some parts of the world (Kager *et al.*, 2007). In 60% of the cases secondary aberrations are found, most often +21, +8, +4, i(7)(q10) and i(9)(q10), but they do not seem to affect prognosis (Pui *et al.*, 1994). The der(19)t(1;19) may be associated with slightly better prognosis than the t(1;19); however, various studies have reported conflicting results (Uckun *et al.*, 1998; Schultz *et al.*, 2007).

The *MLL* gene at 11q23 is highly promiscuous, having several different translocation partners. Aberrations involving *MLL* occur in the majority of children ≤1 year of age affected by ALL and constitute approximately 8% of all pediatric ALL and about 10% of all older ALL patients (Pui *et al.*, 2004). The two most common gene fusions involving *MLL* in ALL are *MLL/AFF1* [t(4;11)(q21;q23)] and *MLL/MLLT1* [t(11;19)(q23;p13)] both associated with

an unfavorable prognosis (Pui *et al.*, 2008). Prognostically insignificant secondary aberrations are found in 25% of the t(4;11) cases (often +X, i(7)(q10) and +8) and in 35% of the t(11;19) cases (often +X and +8) (Moorman *et al.*, 2005).

The t(9;22)(q34;q11) [*BCR/ABL1*] resulting in the Ph characteristic for chronic myeloid leukemia (CML) is found in 2-3% of pediatric ALLs and in about 25% of adult ALLs (Pui *et al.*, 2004). The aberration is associated with very poor prognosis in both children and adults, with allogeneic stem cell transplantation being the only curative treatment (Pui and Evans, 2006). Secondary aberration are found in 50% of the cases, commonly +der(22)t(9;22), -7, +21 and +8. Monosomy of chromosome 7 is associated with an even worse prognosis (Heerema *et al.*, 2004; Li *et al.*, 2009). Two common *BCR/ABL1* transcripts have been described: P190 and P210. The P210 transcript is a hallmark of CML, whereas the P190 is primarily (but not exclusively) associated with ALL (Clark *et al.*, 1987). The P190 transcript is found in nearly 90% of children diagnosed with Ph-positive ALLs (Suryanarayan *et al.*, 1991) whereas adult Ph-positive ALL cases can present with either a P190 or a P210 oncoprotein, or both (Klco *et al.*, 2008).

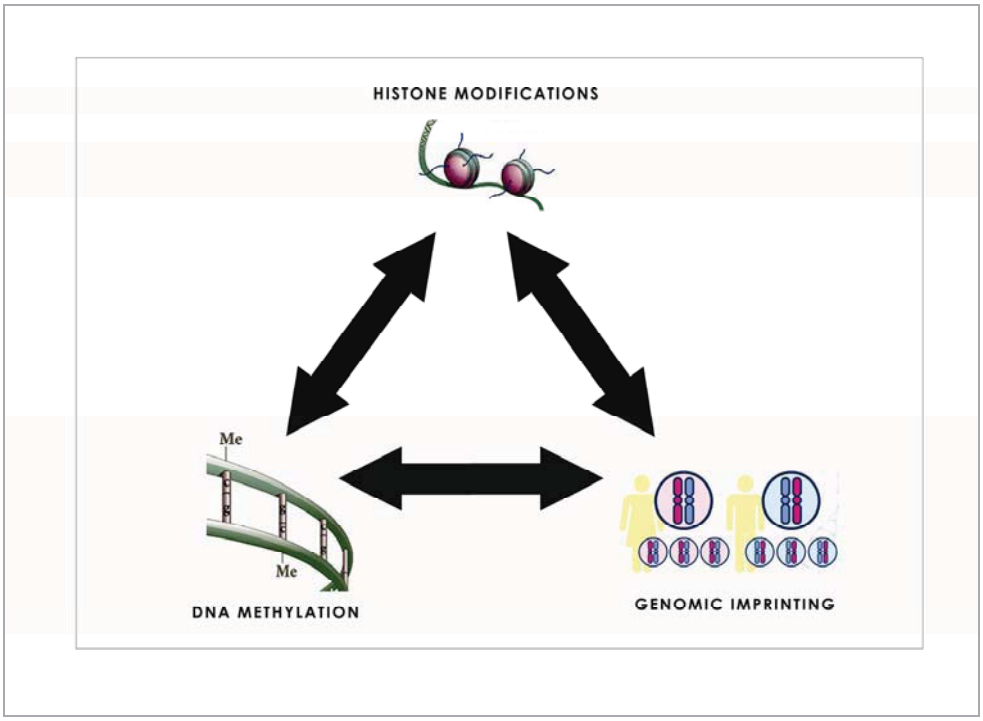


Figure 2. Epigenetic mechanisms. Histone modifications, genomic imprinting and DNA methylation constitute different branches of the field of cancer epigenetics. But in reality these are in fact closely interweaved and simultaneous processes, together interacting and contributing to the regulation of gene transcription.

Epigenetic changes in ALL

The term epigenetics means heritable alterations in gene expression not due to a physical change in the DNA sequence (Bird, 2007). Today, three main types of, closely interacting, epigenetic levels are recognized: histone modifications, genomic imprinting and DNA methylation (**Figure 2)**. During the last two decades, epigenetic alterations have gained recognition as important contributors to tumorigenesis and neoplastic progression (Feinberg and

Tycko, 2004). This may be particularly relevant for ALL, since some translocations arise in utero, but without cooperative secondary genetic or epigenetic events they fail to produce overt leukemia (Greaves, 2005). In addition, since epigenetic events are reversible, they constitute interesting targets for therapeutic intervention, not least in hematologic malignancies (Shaker *et al.*, 2003; Fiskus *et al.*, 2009; Issa and Kantarjian, 2009).

The fact that histone modifications affect transcription has been known for long; however, this epigenetic event was in fact the last one to be linked to cancer. To date, several histone modifications have been described, such as acetylation, methylation, phosphorylation, ubiquination, sumolyation, ribosylation, deamination and proline isomerization. These modifications mainly have two functional outcomes, either they regulate transcription by inducing looser or more tightly packaged chromatin or they serve as docking sites for other chromatin-remodeling protein complexes. A growing number of studies is now emerging, focusing on elucidating the chromatin maps and histone modifications specific for different type of hematologic malignancies (Neff and Armstrong, 2009).

Imprinting, which refers to the selective methylation of maternal and paternal alleles during gametogenesis in such a way that a specific parental allele is exclusively expressed in the embryo (Surani *et al.*, 1984), has also been linked to human neoplasia. Loss of imprinting, leading to aberrant biallelic expression of genes with oncogenic potential, or selective loss of uniquely paternally or maternally expressed tumor suppressor genes has been shown in a number of human neoplasias (Feinberg and Tycko, 2004). In addition, it has been reported that the frequent deletions of chromosome arm 9p seen in ALL leads to preferential loss of maternal alleles, suggesting a germ-line event

promoting leukemogenesis (Morison *et al.*, 2002); however, this has as yet not been confirmed in any subsequent study.

DNA methylation, which is the most well studied branch of epigenetics, involves methyl group addition to the cytosine ring of CpG nucleotides. Although CpG dinucleotides are under-represented in the total genome they are frequently found within the promoter region of human genes. These regions, named promoter-associated CpG islands, usually lack methylation in normal cells; however, this may vary from cell to cell type. This lack of methylation allows transcription to occur (Suzuki and Bird, 2008). Abnormal methylation patterns have been associated with neoplasia for a long time, with cancer cells displaying global hypomethylation and gene promoter-specific CpG hypermethylation. Methylation of promoters is associated with a closed chromatin structure and transcriptional silencing of the associated gene (Feinberg and Vogelstein, 1983). A large number of genes has been implicated as hypermethylated in hematologic malignancies in general (Boultwood and Wainscoat, 2007) and in ALL specifically (Garcia-Manero *et al.*, 2009). Hematopoietic progenitors generally are free from this type of gene silencing (Herman *et al.*, 1997); however, dynamic changes in promoter methylation levels in order to control growth factor and cytokine expression during myeloid development have been demonstrated (Lubbert *et al.*, 1997).

REVIEW OF PEDIATRIC HeH ALL

The HeH subgroup was first described as an entity of its own in the early 1980's when Kaneko and coworkers (1981) reported that modal chromosome numbers ranging from 50 to 59 only occurred in pediatric ALLs and not in adult ALL. Closely thereafter, several independent studies reported the same finding and by the mid 1980's several additional studies had confirmed HeH as a cytogenetic subgroup of childhood ALL (Paulsson and Johansson, 2009). However, it should be stressed that already 1967 Fritz Lampert described that high DNA content was associated with sensitivity to chemotherapy and longer survival in childhood acute leukemias; the favorable prognosis of HeH ALL is now well recognized (see below) (Lampert, 1967).

Cytogenetic features

The HeH subgroup in pediatric ALL is characterized by a nonrandom pattern of gains of chromosomes X, 4, 6, 10, 14, 17, 18 and 21, resulting in modal numbers ranging from 51-67. Chromosome 21 is frequently tetrasomic, even pentasomic in some cases, whereas the remaining commonly gained chromosomes usually are trisomic. Monosomies are rarely seen. In addition to the chromosomal gains, approximately 50% of all HeH cases carry structural chromosome changes identifiable by conventional cytogenetics (Groupe Francais de Cytogénétique Hématologique, 1993; Raimondi *et al.*, 1996;

Forestier *et al.*, 2000a; Moorman *et al.*, 2003; Paulsson and Johansson, 2009). The most common ones are listed below:

About 5% of HeH ALL carries cytogenetically identifiable balanced abnormalities. These are mainly translocations, half of which are common ALL-associated ones, such as the previously mentioned t(1;19) and t(9;22). The other half is non-recurrent balanced rearrangements. To date, no recurring balanced translocation resulting in a novel fusion-gene specific for HeH ALL has been identified (Mitelman *et al.*, 2009).

The most frequent structural changes in HeH are unbalanced, resulting in net gain or loss of chromosomal material. Most common are partial gain of 1q, deletions involving 6q and isochromosome 7q and 17q (Pui *et al.*, 1992; Raimondi *et al.*, 1996; Moorman *et al.*, 2003; Paulsson and Johansson, 2009). Gain of 1q through duplication, isochromosome formation or unbalanced translocations is one of the most frequent acquired genetic abnormalities in neoplasia and has been reported in all major diagnostic subgroups of hematologic malignancies (Mitelman *et al.*, 2009). Despite the high prevalence very little is known about its origin, molecular genetic characteristics and functional outcome. This stimulated us to undertake the first study (*Article I*) of this thesis, a genomic and epigenetic investigation of dup(1q) in pediatric HeH ALL. This work is reviewed later in this thesis. Deletions on chromosome arm 6q occur in about 5% of all HeH cases and are usually independent of dup(1q) (Moorman *et al.*, 2003). Isochromosome 7q are present in about 1-2% of the cases, with the functional outcome being net loss of 7p and 7q gain (Pui *et al.*, 1992; Martineau *et al.*, 1996). Trisomy 17 is one of the characteristic whole chromosomal gains of HeH and 2-5% of all cases harbor i(17q) (Martineau *et al.*, 1996; Raimondi *et al.*, 1996; Moorman *et al.*,

2003). However, gain of chromosome 17 and i(17q) are mutually exclusive indicating that the functional outcome of both aberrations could be gain of 17q (Mitelman *et al.*, 2009; Paulsson and Johansson, 2009).

With the introduction of high resolution copy-number arrays it has become evident that microdeletions are relatively frequent in pediatric HeH ALL. Recurring loss of individual genes such as *CDKN2A*, *ETV6*, *IKZF1*, *PAX5*, *RB1* and *TCF3* results from these submicroscopic aberrations (Mullighan *et al.*, 2007). Compared with other genetic ALL subgroups the overall frequency of microdeletions is lower in HeH; however, this could be explained by potential masking in single nucleotide polymorphism (SNP) arrays of submicroscopic aberrations on trisomic chromosomes by the additional homologue (Paulsson and Johansson, 2009).

Mutations

The RTK-RAS signaling pathway involves, among others, four genes that have been demonstrated to be frequently (~30%) mutated in pediatric HeH ALL: *FLT3*, *KRAS*, *NRAS* and *PTPN11* (Armstrong *et al.*, 2004; Taketani *et al.*, 2004; Tartaglia *et al.*, 2004; Wiemels *et al.*, 2005; Paulsson *et al.*, 2007; Stam *et al.*, 2007). Probably due to the fact that these genes are involved in the same signaling cascade, mutations in these genes are usually mutually exclusive in HeH ALL (Paulsson *et al.*, 2007). In addition, among childhood ALLs harboring *PTPN11* mutations, HeH seem to have a significantly higher prevalence compared to other genetic subtypes, with the mutation being secondary to the HeH karyotype and more common at relapse (Molteni *et al.*, 2009).

Epigenetic changes

The focus of epigenetic studies in childhood ALL has been entirely on delineation of DNA methylation; to date, no studies investigating histone modifications have been published. Furthermore, the DNA methylation studies of HeH ALL have mainly focused on promoter hypermethylation of individual genes. Several tumor suppressor gene promoters have been shown to have an aberrant methylation status, with *CADM1*, *ESR1*, *FHIT*, *RARB*, and *WNT5A* being hypermethylated in more 50% of HeH cases (Paulsson *et al.*, 2009).

To date, we know of only two studies utilizing array-based methodologies to identify differential methylation of a larger number of genes in B-lineage ALL and these studies utilized low-resolution platforms (60 and 2300 genes, respectively) and included mainly adult patients (Scholz *et al.*, 2005; Taylor *et al.*, 2007). Thus, next to nothing is known about the complete DNA methylome of childhood ALL. This led us to investigate methylation on a genome-wide level investigating primary patient samples from t(12;21)-positive and HeH ALL (*Article II*). We analyzed large-scale global methylation profiles of each chromosome using bacterial artificial chromosome (BAC) arrays as well as direct gene silencing through hypermethylation by the use of a microarray platform covering all promoter-associated CpG islands. In addition, global gene expression data were used to investigate correlations between hypermethylated target genes and their expression levels. Furthermore, tiling resolution array comparative genomic hybridization (CGH) was used detect genomic imbalances that might affect gene expression and/or be associated with large scale methylation changes. This work is reviewed in the *The Present Study* section (*Article II*).

Clinical features

Pediatric HeH ALL is associated with a low WBC count; the median in published data is <10 x 10^9/l (Paulsson and Johansson, 2009). Thus, based on WBC count alone these patients are rarely stratified into other than the standard risk group of the Nordic Society of Paediatric Haematology and Oncology (NOPHO) protocol used in the Nordic countries (Schmiegelow *et al.*, 2009). The patients are often moderately anemic and thrombocytopenic at the time of diagnosis; however, a subset displays normal platelet levels.

The immunophenotype of HeH is that of a typical B-cell precursor ALL: $CD10^+$, $CD13^-$, $CD19^+$, $CD22^+$, $CD24^+$, $CD33^-$, $CD34^+$, $CD45^-$, $CD65^-$, $CD66c^+$, $HLA\text{-}DR^+$, cIg^-, sIg^- and TdT^+ (Hrusák and Porwit-MacDonald, 2002). However, the use of immunophenotyping to diagnose HeH is limited, since it cannot distinguish HeH from some other B-cell precursor ALL subtypes. The bone marrow morphology, displaying small and uniformly shaped cells, are more often FAB L1 than L2, and never L3; but the morphology is not used for diagnosis of HeH due to lack of usable characteristics (Smets *et al.*, 1995). Patients with HeH ALL usually harbors a high bone marrow blast percentage (Paulsson and Johansson, 2009).

The outcome of HeH is among the most favorable within ALL, not surpassed by any other cytogenetic subgroup (Kaneko *et al.*, 1981; Pui and Evans, 2006). Initial treatment failure is very rare and subsequently complete morphological remission is reached in the great majority of cases. The event-free survival is above 80% and the overall survival currently reaches approximately 90% (Forestier *et al.*, 2008a; Pui *et al.*, 2008). It is therefore apparent that treatment reaches a second successful remission in many of the relapsed patients.

Incidence

HeH are found in about 25% of all pediatric ALLs, but in less than 10% of adult ALLs (Mitelman *et al.*, 2009). This makes it the most common ALL subgroup in children. All major studies of pediatric HeH ALL, derived from different parts of the world, have reported a prevalence around 25-30%, indicating that there is no major geographic incidence heterogeneity (Paulsson and Johansson, 2009). However, it should be noted that most studies did not include patients from developing countries. Whether the abovementioned frequency is influenced by socioeconomic factors thus remains unknown. Ethnic background could have an influence HeH ALL. It has been reported that HeH ALL is underrepresented among African-Americans, compared to other ethnic groups (Pui *et al.*, 1995; Raimondi *et al.*, 1996; Pollock *et al.*, 2000; Bhatia *et al.*, 2002). Whether this is due to genetic or environmental factors is unknown; perhaps genome-wide association (GWA) studies of this ALL subtype in different ethnic groups could shed some light on this issue.

HeH displays a pronounced and very characteristic age peak at 2-4 years of age, being very rare in infant leukemia and uncommon above the age of seven years. HeH together with t(12;21)-positive ALL comprise about 80% of all ALLs occurring in the 2-7 year age span (Forestier and Schmiegelow, 2006). The reasons for this specific age peak of the two most common forms of pediatric ALLs are discussed below, in relation to HeH leukemogenesis. The ratio of male/female incidence based on all cases of HeH ALL below the age of 18 years present in the Mitelman Database of Chromosome Aberrations in Cancer (Mitelman *et al.*, 2009) is 1.3, revealing a slight male preponderance. However larger individual series have demonstrated a sex ratio close to 1.0,

making it uncertain whether HeH is really more common in boys (Paulsson and Johansson, 2009).

Etiology

Considering that HeH is the most common subgroup in pediatric ALL our knowledge about its etiology is surprisingly poor. Looking at the spectrum of genetic risk factors, several Mendelian disorders have been associated with increased risk of developing acute leukemias; however, not specifically HeH ALL (Benson and Horwitz, 2006). Among constitutional genetic conditions, only Down syndrome has been associated with an elevated risk for developing HeH ALL (Hasle, 2001; Benson and Horwitz, 2006). This constitutional chromosome disorder is associated with an overall 20-fold higher risk for developing ALL, but the proportion of HeH ALL is lower in this patient population compared to non-Down syndrome pediatric patients with ALL. However, the total risk of developing HeH ALL is still higher with, than without, a constitutional trisomy 21 (Forestier *et al.*, 2008b).

Concerning individual genes, the immune response gene *HLA-DBP1*0201* has been demonstrated to occur more frequently in pediatric patients with HeH ALL compared to healthy controls (Taylor *et al.*, 2002; Taylor *et al.*, 2008; Taylor *et al.*, 2009). This could indicate that it is associated with an increased risk; however, no follow-up studies from other groups have as yet confirmed these results. In two recent independent GWA studies, three SNPs in the gene *ARID5B* on 10q21.2 were associated with increased risk for HeH ALL compared to healthy controls (Papaemmanuil *et al.*, 2009; Treviño *et al.*, 2009). Two of these SNP were demonstrated to be linked with methotrexate polyglutamate accumulation in bone marrow (Treviño *et al.*,

2009), a clinical feature clearly linked to HeH ALL (Synold *et al.*, 1994; Kaspers *et al.*, 1995; Zhang *et al.*, 1998; Belkov *et al.*, 1999; Kager *et al.*, 2005; Whitehead *et al.*, 2005). These germline variations in the *ARID5B* locus might affect the gene's function in B-lineage development and create a susceptibility to HeH ALL, but the mechanisms remain unknown. As regard genetic risk factors, it should however be noted that studies of siblings to HeH patients have shown that they do not have an increased risk of developing ALL, indicating that constitutional genetics has a limited impact on the etiology of pediatric HeH ALL (Schmiegelow and Hjalgrim, 2006).

Little is known about environmental risk factors for HeH ALL. It has been reported that pediatric HeH ALL is less common in children exposed to paternal smoking, something that the authors speculated could be due to that hyperdiploid cells might be sensitive to toxic smoking-associated metabolites (Wiemels *et al.*, 2005). Elevated birth weight and low maternal folate consumption during pregnancy have also been associated with increased general risk of pediatric ALL (Thompson *et al.*, 2001; Hjalgrim *et al.*, 2003b). That birth weight is as a risk factor for HeH ALL has been shown (Hjalgrim *et al.*, 2004), whereas the risk-association with low folate intake did not pinpoint specific genetic subgroups. Interestingly, folate deficiency has been shown to induce aneuploidy of chromosomes 17 and 21 in cultured lymphocytes, suggesting a link between hyperdiploidy and folic acid (Wang *et al.*, 2004). In addition, specific polymorphisms of the *MTHFR* gene, involved in folate metabolism, have been demonstrated to be underrepresented in HeH ALL (Wiemels *et al.*, 2001), further supporting the involvement of folic acid in pediatric HeH ALL etiology.

Origin of HeH

How does a cell become HeH? To answer this biologically, and perhaps clinically, important question, four different mechanisms have been suggested: (1) development of a near-haploid cell followed by chromosome-doubling, (2) tetraploidization of the cell followed by chromosome losses, (3) sequential gain of individual chromosomes through several consecutive independent cell divisions and (4) simultaneous gain of chromosomes in one single abnormal mitosis (Onodera *et al.*, 1992). By demonstrating a 2:2 allelic ratio and specific tetraploid pattern of whole chromosome uniparental isodisomies (UPDs), or no pattern of UPDs at all, in virtually all cases investigated, HeH seems to arise by the simultaneous gain mechanism in about 70% of the cases and by the tetraploid pathway in the remaining 30% of the cases (Paulsson *et al.*, 2003; Paulsson *et al.*, 2005; Paulsson and Johansson, 2009).

The accepted theory that hematologic malignancies are initiated and maintained by leukemic stem cells have been supported by the detection of gene fusions, such as [*ETV6/RUNX1*] caused by the t(12;21), in specific precursor lineages (Castor *et al.*, 2005; Hong *et al.*, 2008; Bernt and Armstrong, 2009). However, few studies have addressed this issue in HeH ALL and the results have also been ambiguous. Studies by Larramendy *et al.* (1995) and Quijano *et al.* (1997) concluded that pediatric HeH ALL might well be a disease originating from a pluripotent stem cell since hyperdiploidy were found in $CD34^{+}CD33^{-}CD38^{-}CD19^{-}$ cells deemed to belong to primitive compartments of the bone marrow, whereas Kasprzyk *et al.* (1999) detected trisomies only in commited B cell progenitors; however, they could not exclude a stem cell involvement. Le Viseur *et al.* (2008) could by engraftment of flow sorted blast derived from two HeH ALL patients demonstrate that at

least some cases of HeH have leukemia-initiating activity in populations consistent with later mid-stage development. So, in conclusion, this issue still remains unresolved.

Leukemogenesis

HeH is very likely of utmost importance for leukemogenesis; however, there is ample evidence that it in itself it is not sufficient for leukemic transformation. The strongest support in favor of this has come from studies of neonatal blood spots identifying preleukemic cells carrying clonotypic *IGH@* rearrangements identical to those later detected in the neoplastic HeH cells (Yagi *et al.*, 2000; Panzer-Grümayer *et al.*, 2002; Taub *et al.*, 2002; Maia *et al.*, 2004). In addition, HeH cells with clonotypic *IGH@* rearrangements have been detected in a pair of monozygotic twins with concordant pediatric HeH ALL, where it most likely reflects blood cell chimerism resulting from shared placental vasculature, *i.e.*, the leukemia starts in one twin in utero and spreads to the other via the placenta (Maia *et al.*, 2003). Thus, it could be argued that hyperdiploidy most likely occurs prenatally but secondary events are needed for overt disease. This is further strengthen by the typical age peak of diagnosis around 2-4 years of age shared with the t(12;21) leukemia, where identical *ETV6/RUNX1* genomic fusions have been demonstrated to occur both in preleukemic cells derived from neonatal blood spots and in later diagnostic samples (Greaves, 2005). The latency of disease both for t(12;21)-positive and HeH ALL could, if it is of prenatal origin, reflect the time span needed to acquire additional aberrations necessary for overt disease. Indeed, accompanying genetic lesions are found in a majority of HeH cases (Mitelman *et al.*, 2009). However, the temporal order of these additional aberrations

remains to be ascertained since it has not yet been clearly demonstrated that they indeed occur after the HeH pattern.

Gene dosage

What is the functional outcome of HeH? Since tetrasomies have been demonstrated to display equal duplication of parental homologues (2:2 allelic ratio) in the majority of cases, neither imprinting defects nor selective duplication of mutated genes seems to be a valid explanation for the pathogenetic mechanism associated with HeH and neoplastic transformation (Paulsson *et al.*, 2003; Paulsson *et al.*, 2005; Paulsson and Johansson, 2009). Expression dysregulation due to gene dosage effects stemming from the massive but nonrandom aneuploidy is a more attractive hypothesis. Indeed, it has been demonstrated that pediatric HeH ALL has an expression signature that is distinctly separate from other pediatric ALL subtypes and that the great majority of genes on the gained chromosomes in HeH display upregulation compared to other ALL subgroups without tri- or tetrasomies (Andersson *et al.*, 2005). In addition, we recently demonstrated global overexpression of genes on the additional chromosomes also when comparing disomic and trisomic/tetrasomic chromosomes within the HeH group itself (*Article II*) (**Figure 3**).

Relapse

Despite the generally very good prognosis of pediatric HeH ALL, approximately 20% of the cases relapse and about 10% of all patients succumb to the disease. The reasons for relapse remain unknown. One possibility could be that the genetic features at the time of diagnosis differ

between cases that subsequently relapse and those who remain in long-term remission. In this respect it has been suggested that HeH with structural genetic changes, in particular i(17q), are associated with a poorer outcome (Pui *et al.*, 1988; Pui *et al.*, 1989); however, this has not been confirmed in later studies (Raimondi *et al.*, 1996; Forestier *et al.*, 2000b; Moorman *et al.*, 2003). Furthermore, some investigators have reported a superior outcome for cases harboring certain trisomies, that is +4, +10,+17 and +18, findings, at least as regards +4, +10 and +17, that are now used for risk stratification by the Children's Oncology Group (COG) (Schultz *et al.*, 2007).

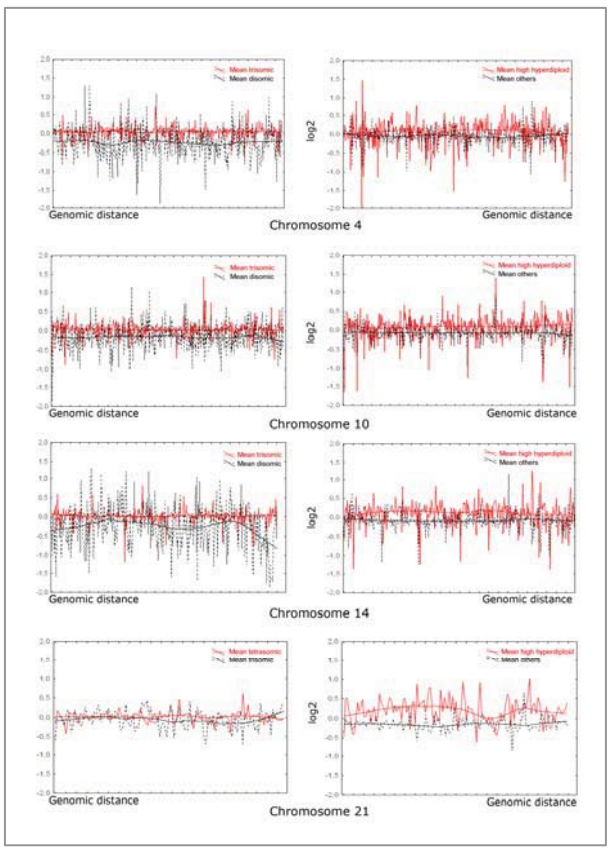


Figure 3. Expression patterns of HeH ALL. Lowess curves of mean-centered expression ratios of disomies, trisomies and tetrasomies for chromosomes 4, 10, 14 and 21 among the HeH cases (left) and between HeH cases and all other pediatric B-lineage ALLs (right) (*Article II*).

Whether the patterns of micro-deletions or mutated genes that recently have been detected in pediatric HeH ALL (Paulsson *et al.*, 2006; Kuiper *et al.*, 2007; Mullighan *et al.*, 2007; Paulsson *et al.*, 2008) have a prognostic impact is currently unknown. One approach to identify changes possibly associated with relapse is to investigate paired diagnostic and relapsed samples. This has to date only been performed in two SNP array studies of a total of eight cases of HeH ALL (Mullighan *et al.*, 2008; Yang *et al.*, 2008). These studies report contradictory findings as regards genetic heterogeneity/homogeneity between pediatric HeH ALL at diagnosis and relapse: the two patients analyzed by Yang *et al.* (2008) displayed identical numbers of aberrations at diagnosis and relapse, whereas Mullighan *et al.* (2008) observed differences in copy number alterations, involving both whole chromosomes and/or micro-aberrations, in diagnostic and relapse samples in five of the six cases analyzed. Clearly further studies are needed to clarify this important issue. To address this, we performed high resolution SNP array and mutation analyses of 11 paired samples from pediatric HeH ALL (*Article III*).

THE PRESENT STUDY

This section includes the aims of thesis, a brief description of the material and methods used, as well as summary of the main results of each individual article. At the end, a general discussion is given together with the major conclusions of this thesis, and finally some remarks on future directions. For a more thorough description of the individual studies, the reader is referred to the original articles (*Articles I-III*).

Aims

The general aim of this thesis has been to broadly characterize pediatric HeH ALL using a wide spectrum of genomic and epigenomic techniques in order to better understand the pathogenetic mechanisms underlying leukemogenesis, disease progression and potential relapse of this leukemic subtype. More specifically the aims were:

- to perform a genomic and epigenetic investigation of the most common structural chromosome change in pediatric HeH ALL and Burkitt lymphoma (BL), namely dup(1q) (*Article I*).

- to perform genome-wide methylation profiling coupled to global gene expression and copy-number arrays on pediatric HeH ALL, using t(12;21)-positive ALL as a comparison, in order to investigate if

aberrant methylation patterns play a significant leukemogenic role (*Article II*).

– to identify genetic changes associated with relapse and to gain information about the temporal order of genetic abnormalities and clonality by investigating paired diagnostic and relapse samples from pediatric HeH ALL patients, using a genome-wide screening method as well as mutation analyses of RTK-RAS genes (*Article III*).

Patients and samples

Article I

This study comprised six and four cases of dup(1q)-positive HeH ALL and BL, respectively. All HeH ALL patients had been diagnosed and treated according to the NOPHO-ALL 1992 or 2000 protocols at the Departments of Pediatric Oncology and Hematology, Lund University Hospital, Lund, and Linköping University Hospital, Linköping, Sweden. Bone marrow and/or peripheral blood were collected at diagnosis. The breakpoints of the dup(1q) were investigated in all cases. In addition satellite II methylation was investigated in seven of the ten cases, as well as in four and three non-dup(1q) HeH ALLs and BLs, respectively, using normal peripheral blood and ICF B-lymphocytes as controls in each assay. Finally, gene expression data from three of the ten cases were retrieved from a previously reported dataset (Andersson *et al.*, 2005)

Article II

Samples from 20 children with B-cell precursor ALL, comprising ten with HeH and ten with t(12;21)(p13;q22), were included in the study. The patients had been diagnosed and treated according to the NOPHO-ALL 1992 or 2000 protocols at the Departments of Pediatric Oncology and Hematology, Lund University Hospital, Lund, and Linköping University Hospital, Linköping, Sweden. Bone marrow (n=19) and peripheral blood (n=1) samples were collected at diagnosis and the samples were cytogenetically characterized by conventional chromosome banding analysis. The presence of the cytogenetically cryptic [*ETV6/RUNX1*] was identified by RT-PCR and verified by FISH, as part of routine diagnostic analyses. The additional chromosomes identified by G-banding in the HeH cases were all confirmed by array CGH.

Article III

Diagnostic and relapse samples from 11 children with pediatric HeH ALL were included. The patients had been diagnosed and treated according to the NOPHO-ALL 1992 or 2000 protocols at the Departments of Pediatric Oncology and Hematology, Lund University Hospital, Lund; Karolinska Institute, Stockholm; Linköping University Hospital, Linköping; Umeå University Hospital, Umeå; Uppsala University Hospital, Uppsala, Sweden; and Rigshospitalet, Copenhagen, Denmark. Cytogenetics of bone marrow samples obtained at diagnosis and relapse were performed using conventional methods. None of the cases had the well-known ALL-associated translocations t(1;19)(q23;p13), t(9;22)(q34;q11), 11q23/*MLL* rearrangements or t(12;21)(p13;q22).

Methods

Article I

The proximal as well as the distal breakpoints of the dup(1q) were analyzed using tiling resolution array CGH, which is a method where test and reference DNA are differentially labeled with fluorochromes and subsequently hybridized to slides containing defined genomic sequences. This enables detection of genomic copy-number loss or gain (Pinkel and Albertson, 2005). The 32 k slides used in our study contained 32 433 tiling BAC clones, covering at least 98% of the human genome, and were produced at the SCIBLU DNA microarray resource center at Lund University, Sweden. Labeling of DNA, slide preparation and hybridization were performed as described previously (Jönsson *et al.*, 2007). Initial analyses of the microarray images were performed using the GenePix Pro 4.0 software (Axon Instruments, Foster City, CA). All additional analyses were performed in the Bioarray software environment (BASE) database (Saal *et al.*, 2002) and data normalization was performed for each array subgrid using a lowess curve fitting (Yang *et al.*, 2002). Classification as gain or loss was based on identification as such by the BASE CGH plotter and also by visual inspection of the log2 ratios.

Satellite II domain methylation status was investigated by Southern blot analysis as previously described (Wong *et al.*, 2001). In short, test DNA was digested using a CpG methyl sensitive restriction enzyme and then blotted to a membrane where it was probed with a radiolabeled oligonucleotide. The amount of fractioned DNA reflects the methylation status of the region investigated (Wong *et al.*, 2001). DNA samples were digested using *Bst*BI (New England BioLabs, Ipswich, MA), and the satellite II probes used were an 18 mer single-stranded oligonucleotide and a cloned Chr1-specific insert

excised from the recombinant plasmid pUC1.77 (Cooke and Hindley, 1979). In each Southern blot analysis, normal peripheral blood and ICF B-lymphocyte (Corriell Cell Repositories, Camden, NJ) (Jeanpierre *et al.*, 1993) DNA was included as methylated and hypomethylated controls, respectively. Using a FLA 3000 phosphoimager (Fujifilm Corporation, Tokyo, Japan), the approximate amount of methylation was assessed by comparing the intensities of the hybridized fragments.

The cDNA microarray analyses had previously been reported (Andersson *et al.*, 2005). The methodology is similar to that of array CGH, but instead of quantifying gene copy-numbers, the hybridization reflects RNA levels and hence gene expression. Samples were hybridized to 27 k microarray slides containing 25 648 cDNA clones (SCIBLU DNA Microarray Resource Center) representing 13 737 Unigene clusters and 11 592 Entrez gene entries, according to the Unigene build 195. RNA extraction, amplification, labeling, hybridization, scanning, post-hybridization washing and feature analysis were performed as described by Andersson *et al.* (2005). The data analyses were all performed in BASE. To identify differentially expressed genes in the dup(1q)-positive ALLs, a t-test comparing HeH ALLs with and without gain of 1q was performed on the expression patterns of genes in the minimally gained 1q region.

Article II

The global methylation levels were measured using methylated DNA immunoprecipitation (MeDIP) (Weber *et al.*, 2005) followed by hybridization to two different microarray systems (**Figure 4**). For genome-wide profiling we used tiling BAC arrays in order to obtain large-scale methylation profiles of

each chromosome and for gene promoter-specific analyses we used a microarray platform (Roche NimbleGen, Madison, WI) comprising a total of 28,226 CpG islands.

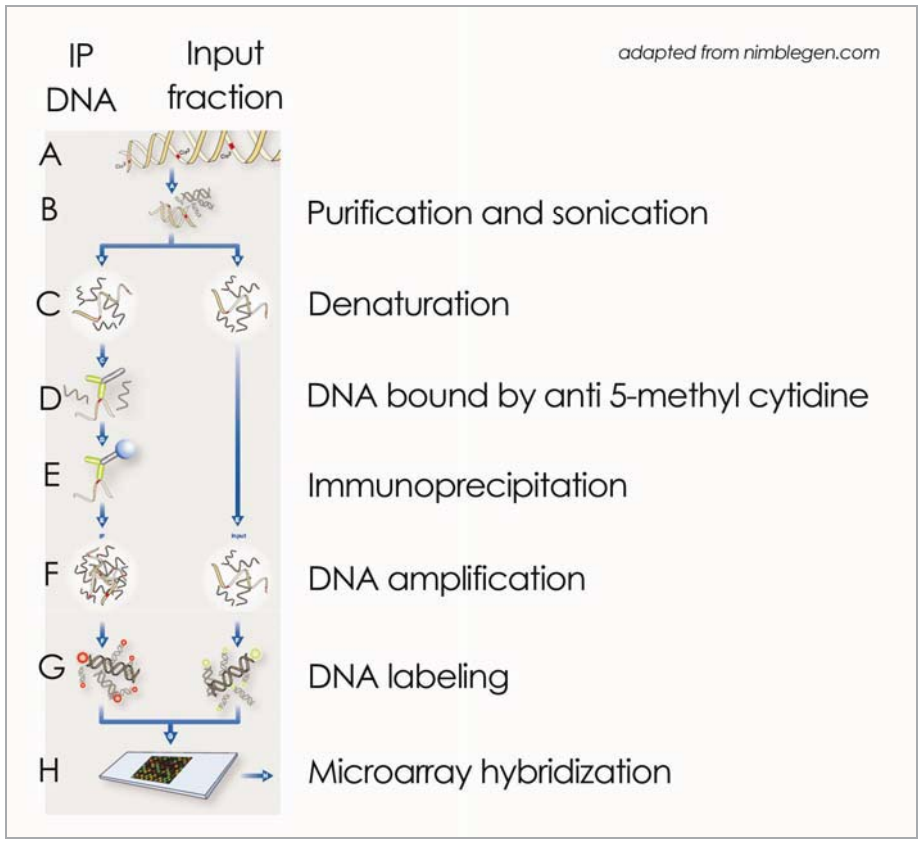


Figure 4. MeDIP. A) Selection of test DNA. B) High grade purification and fragmentation by sonication. C) DNA denaturation and input fraction (IF) collection. The IF serves as reference in the subsequent microarray hybridizations. D) DNA is treated with a monoclonal antibody against 5-methylcytidine. E) Immunoprecipitation using Protein A agarose beads and subsequent DNA recovery. F) Immunoprecipitated DNA (IP DNA) and IF is amplified. G) IP DNA and IF are differentially labeled using fluorochromes. H) Mixing and hybridization to a suitable microarray slide.

Unsupervised clustering and principal component analyses were used to investigate methylation differences/similarities between the BAC array large-scale methylation profiles derived from t(12;21)-positive and HeH cases. The two groups were also analyzed in relation to methylation and gene density on each BAC clone to detect crude methylation differences between individual chromosomes in each subtype. The detected highly methylated genes in the promoter and CpG island array were analyzed in three different ways: First, all genes with significantly high methylation levels were grouped according to the genetic ALL subtype in order to identify those that were uniquely hypermethylated in either t(12;21)-positive or HeH cases. Second, an edge preserving smoother analysis determining a sequence with as few jumps as possible was performed. This segmentation analysis used the Potts filter with a penalty parameter in order to identify highly methylated genes throughout the genome (Lingjaerde *et al.*, 2005). Third, we identified all genes that displayed a high level of methylation in ≥30% of cases in either genetic ALL subgroup or that were present in a dataset comprising genes known to be hypermethylated in malignant disorders, involved in B-cell development and mutated or rearranged in B-lineage ALL. Each gene was then functionally annotated using the EASE software (Hosack *et al.*, 2003) and correlated with the gene expression data described below.

Identification of methylation hotspots, *i.e.*, genomic sites harboring a significantly higher number of highly methylated genes compared with the rest of the genome, was performed by plotting the spatial and number distribution of promoters enriched for methylation using a density estimator (Lingjaerde *et al.*, 2005). In order to designate a chromosomal segment as a methylation hotspot four probes had to be present within the segment analyzed.

To test if methylation hotspots were associated with imprinted genes we used the mapping information provided by Luedi *et al.* (2007) on all known imprinted human genes. We ascertained the number of imprinted genes per chromosome band. Then, we calculated how often bands with a specific number of imprinted genes harbored methylation hotspots.

To validate the reliability of the promoter array, eight hypermethylated genes were selected for bisulfite sequencing (**Figure 5**). The bisulfite-treated DNA was used as template in a standard PCR amplification utilizing primers directed towards the promoter regions harboring the hypermethylated CpG islands, as indicated by the array. The primers were designed using the MethPrimer software (Li and Dahiya, 2002), the PCR products sub-cloned using the TOPO-TA system (Invitrogen, Carlsbad, CA) and sequenced using standard methods. Finally, the sequences were analyzed using the BiQ Analyzer software (Bock *et al.*, 2005).

Gene expression data based on cDNA microarrays (Andersson *et al.*, 2005) were used to investigate correlations between hypermethylated target genes and their expression levels. Furthermore, array CGH (Jönsson *et al.*, 2007) was performed on each case to detect genomic imbalances that might affect gene expression and/or be associated with large scale methylation changes.

Article III

SNP array analyses were performed using the Affymetrix GeneChip Human 250k *Nsp*, 250k *Sty*, and 10k 2.0 array systems (**Figure 6**) in order to identify genomic copy-number alterations (Heinrichs and Look, 2007). These chips cover approximately 510,000 SNPs, with a median physical distance between

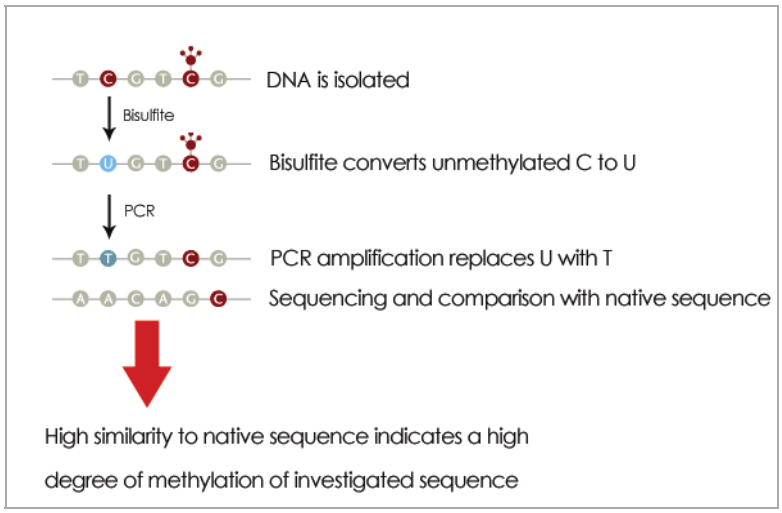


Figure 5. Bisulfite sequencing. DNA is treated with sodium bisulfite which converts cytosine CpG motif residues into uracil, unless a methyl group is bound to the carbon 5 of the cytosine. Using PCR amplification, uracils are converted to thymines in the DNA segment, which subsequently can be sequenced and compared with the native sequence. Conserved cytosines indicate methylation whereas replacement by thymine means that the base is unmethylated.

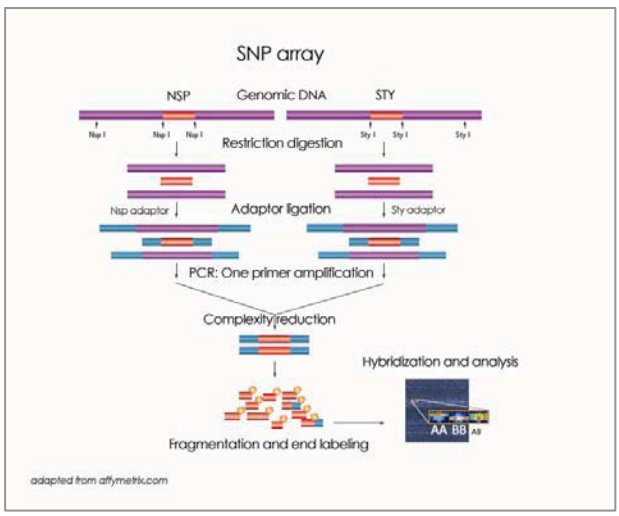


Figure 6. Affymetrix GeneChip. DNA is digested with *Nsp* and *Sty* and ligated to adaptors. An adaptor-specific PCR assay is used to amplify fragments. Products from each restriction digest are combined. The amplified DNA is then fragmented, labeled and hybridized to a SNP array.

the SNPs of <2.5 kb. Hybridization and washes were performed as described previously (Paulsson *et al.*, 2008). The Affymetrix GTYPE software was used for analysis of signal intensity and for genotype calling. As an initial step to identify putative genomic imbalances, log2 ratios were segmented with the CBS algorithm (Venkatraman and Olshen, 2007) using the DNA copy package in R (http://www.bioconductor.org). For further copy number and UPD analyses, the in-house Genome Orientated Laboratory File (GOLF) and the dChip (Lin *et al.*, 2004) software were used. Classification as a copy number alteration was based on visual inspection of the inferred log2 ratio versus the pooled signal intensity of 10 control samples.

For analyses of RTK-RAS genes, codons 835 and 836 in the second tyrosine kinase domain and of internal tandem duplication of exons 14 and 15 in *FLT3*, codons 12, 13 and 61 in *NRAS* and *KRAS* and of exons 3 and 13 in *PTPN11* were sequenced, as described previously (Paulsson *et al.*, 2007). The PCR products sense and the antisense strands were directly sequenced and analyzed using the Seqscape software (PE Applied Biosystems, Foster City, CA).

Results

Article I

The proximal breakpoints in the ten dup(1q)-positive BLs and HeH ALLs were all near-centromeric, with eight of them clustering within a 1.4 Mb segment in 1q12-21.1. The distal 1q breakpoints were more heterogeneous, being more distal in the HeH ALLs than in the BLs. The minimally gained 1q segments in the ALLs and BLs were 57.4 Mb [dup(1)(q22q32.3)] and 35 Mb [dup(1)(q12q25.2)], respectively. Satellite II DNA on 1q was not

hypomethylated, as ascertained by Southern blot analyses of 15 BLs/ALLs with and without gain of 1q, indicating that aberrant methylation was not involved in the origin of dup(1q). Global gene expression analyses revealed that 5 genes in the minimally Mb gained region – *B4GALT3*, *DAP3*, *RGS16*, *TMEM183A* and *UCK2* – were significantly overexpressed in dup(1q)-positive ALLs compared to non-dup(1q) HeH ALL. The *DAP3* and *UCK2* genes were among the most overexpressed genes in the sole dup(1q)-positive BL case investigated.

Article II

Unsupervised clustering and principal component analyses of the BAC array chromosome-wide data successfully subgrouped the t(12;21)-positive and HeH cases, indicating that they are characterized by different methylomes (**Figure 7A**). When comparing the mean methylation levels in the two ALL types, differences between the HeH and the t(12;21)-positive ALLs as regards chromosome-wide methylation on chromosomes 6, 10, 14, 16, 17, 18, 19 and 21 were clearly seen, with the former cases generally displaying a lower methylation of these chromosomes (**Figure 7B**). The majority of these chromosomes are the ones commonly gained in HeH ALL. Furthermore, when comparing the mean methylation levels within HeH ALL, they were generally lower in cases with trisomy/tetrasomy than in cases with disomies of the same chromosomes. Thus, gains in HeH cases are associated with decreased methylation of additional chromosomes. This was not seen in t(12;21)-positive ALL.

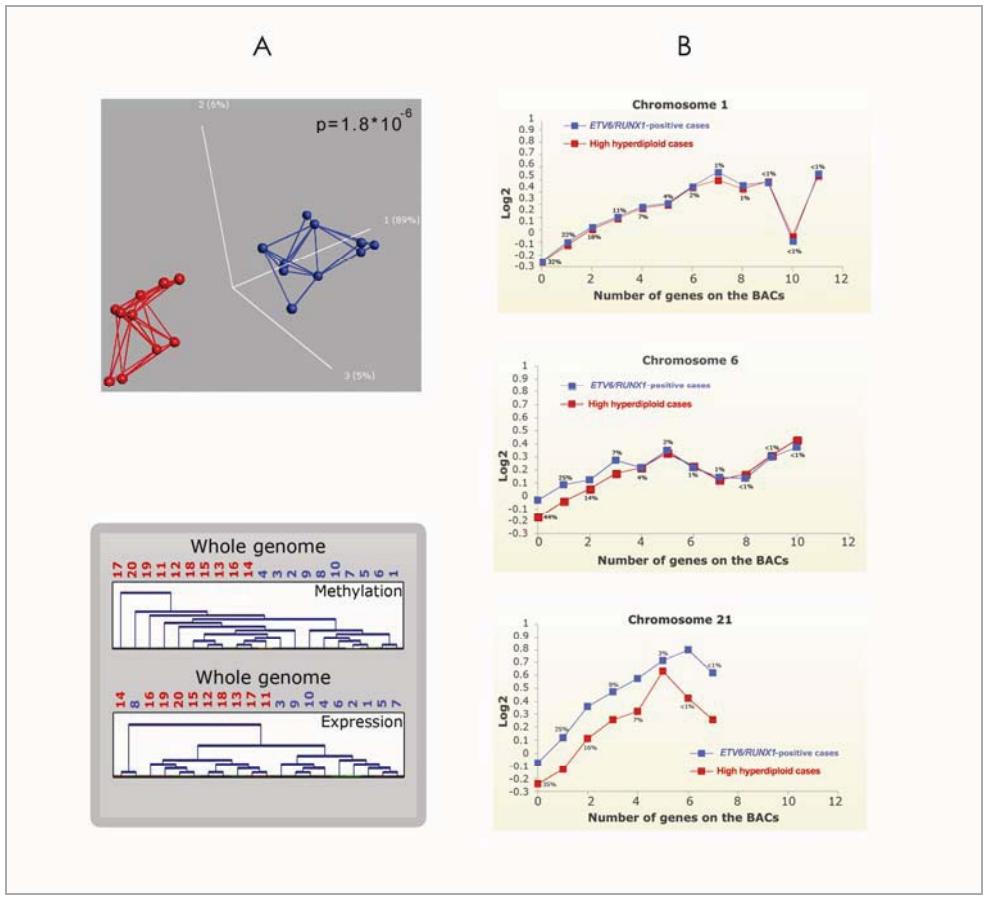


Figure 7. BAC array chromosome-wide methylation profiling A) Principal component and unsupervised clustering analyses of the data could successfully subgroup the majority of t(12;21)-positive (red) and HeH ALLs (blue). B) Mean log2 ratios of the HeH (red) and t(12;21)-positive (blue) ALLs plotted against genomic positions on chromosomes 1, 6 and 21. Similar methylation levels are seen for chromosome 1, which is disomic in both subtypes. Lower methylation of gained chromosomes is seen for both chromosomes 6 and 21 in HeH, in which they were gained.

A total of 8,662 genes with significant methylation scores were identified by the CpG promoter array. The HeH ALLs harbored more hypermethylated genes (n=7650) than the t(12;21)-positive cases (n=3983) and the genes were distributed on all chromosomes, without any clear association with chromosomes frequently gained or lost in either subgroup. Segmentation analysis identified 138 genes that were highly enriched for methylation. It was apparent that enrichment for methylation not necessarily resulted in decreased expression. However, the majority of hypermethylated (60%) genes did show a significantly lowered expression. In addition, a total of 167 different recurring genes and/or genes involved in B-cell development or mutated or rearranged in B-lineage ALL were found to be targets for hypermethylation; 34% of these genes showed a significant decrease in expression.

A total of 58 different methylation hotspots were identified, with 30 of these being common to both t(12;21)-positive and HeH ALLs. The t(12;21)-positive ALLs harbored eight unique methylation hotspots, whereas the HeH ALLs had 20 unique methylation hotspots. In addition, we detected a correlation between the presence of imprinted genes and methylation hotspots, with the frequencies of methylation hotspots increasing with the number of imprinted genes.

In the validation assay of the promoter CpG island methylation profiling, the bisulfite sequencing results corresponded to the array results in all cases, *i.e.*, low or moderate methylation was found when the investigated regions were situated outside the peak and hypermethylation was detected when the analyzed segments were situated within the peak region.

Article III

When analyzing and comparing the 11 investigated paired diagnostic and relapse samples in HeH ALL, the following gains, in decreasing frequency order, were detected: +6 (100%), +21 (100%), +4 (91%), +17 (91%), +18 (91%), +X (86%), +10 (82%), +8 (77%), +14 (45%), +5 (23%), +2 (9%), +7 (9%), +9 (9%), +15 (9%) and +16 (5%). Only one sample harbored a monosomy (-19), and this was not present at relapse. In seven patients, the gains were identical at diagnosis and relapse, whereas they differed by 1-5 trisomies/tetrasomies in four cases. Simultaneous trisomies of chromosomes 4, 10, and 17 – the "triple trisomies" used for stratification of patients into lower-risk by the COG (Schultz *et al.*, 2007) – were found in seven (64%) of the 11 patients.

The number of structural changes varied between 0 and 6 among the 22 samples with a total of 34 different hemizygous deletions, two homozygous deletions and five copy-number gains being identified. Among all 41 structural changes detected, nine were >10 Mb in size; the remaining 32 imbalances may be considered cytogenetically cryptic. None of the five copy-number gains were recurrent. Among the 36 hemi- and homozygous deletions identified, four were recurrent: 7p12.2/*IKZF1*, 9p13.2/*PAX5*, 9p21.3/*CDKN2A* and 9p24.1/*AK3*. None of the homozygous deletions or copy-number gains was shared in paired diagnostic and relapse samples and apart from a homozygous 9p deletion at diagnosis in case 2, these changes were only found at relapse. Among the 34 hemizygous deletions, nine were seen both at diagnosis and relapse in five patients. Of the remaining 25 hemizygous deletions 17 were found at relapse and 8 at diagnosis. Taken together, the mean frequency of structural changes at diagnosis was 1.6, whereas the corresponding frequency

at relapse was 2.9, demonstrating a significantly ($p<0.05$) higher frequency of such aberrations at relapse. UPDs were only detected in two of the 11 paired samples, comprising a UPD8 at diagnosis and relapse. Five of the 11 paired samples harbored a total of six mutated RTK-RAS genes: one *FLT3* mutation, two *KRAS* mutations and three *PTPN11* mutations. Two different mutations, one in *FLT3* and one in *PTPN11*, were found in one single case. Among all mutations, only one was detected both at diagnosis and relapse. Of the remaining five mutations, one was found at diagnosis and four at relapse.

Based on the clonal relationship between the diagnostic and relapse samples three distinct groups were identified: 1) identical genetic changes at diagnosis and relapse (18%), 2) clonal evolution with all changes at diagnosis being present at relapse (18%), and 3) ancestral clones with some changes present at diagnosis but not at relapse (64%), suggesting the presence of an preleukemic clone (Mullighan *et al.*, 2008) (**Figure 8**).

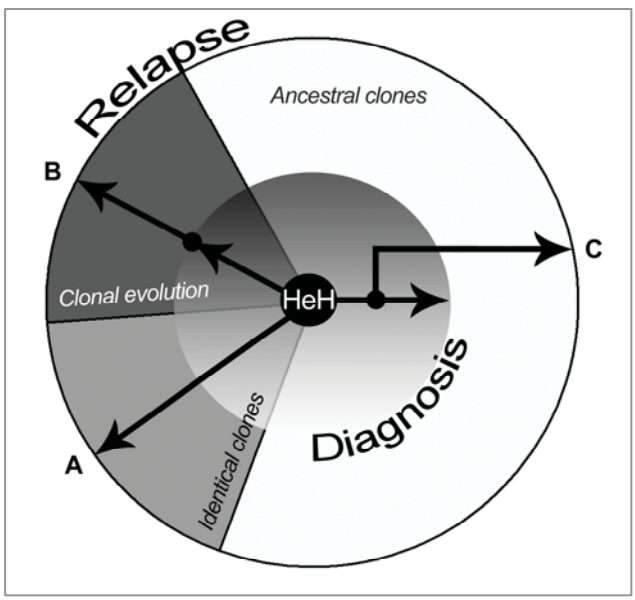


Figure 8. The clonal relationship of HeH ALL were divided into three groups. A) Identical clones at diagnosis and relapse. B) Clonal evolution with all changes at diagnosis being present at relapse. C) Ancestral clones, *i.e.* some changes were conserved and others lost or gained in the diagnostic and relapse samples, respectively.

Discussion

Article I

The results in *Article I* demonstrated near-centromeric proximal 1q breakpoints in all BLs and HeH ALLs investigated. Together with data, derived from other high resolution array studies (Paulsson *et al.*, 2006; Mullighan *et al.*, 2007), this strongly indicates the presence of a breakprone region near the centromeric region of 1q. Why this region is prone to break and to form duplications is unknown; however, it has been hypothesized that hypomethylation of pericentric heterochromatin could be associated with decondensation and instability of satellite sequences, resulting in such rearrangements (Sawyer *et al.*, 1998). However, we found no hypomethylation of the 1q satellite II domain in our series. It is thus highly unlikely that aberrant methylation of the satellite II region should play an important role in the formation of dup(1q), at least not in BLs and HeH ALLs. A more likely explanation could be that the genomic architecture as such, in particular segmental duplications, could be important for this phenomenon. The presence of low-copy repeats has been implicated in the formation of both constitutional and neoplasia-associated chromosomal abnormalities (Ji *et al.*, 2000). Interestingly, such duplications can be found immediately distal to the highly repetitive satellite DNA motifs, such as satellite II, located close to the centromere (Horvath *et al.*, 2001). In addition, small segmental duplications have been mapped to the region on 1q in which the proximal breakpoints occurred in our cases (Sharp *et al.*, 2006). The frequent occurrence of dup(1q) could therefore be explained as a consequence of these rearrangement hotspots.

Does dup(1q) play any role in leukemogenesis? Murine models inoculated with human leukemic B-cell clones carrying dup(1)(q11q32) developed tumors that were more tumorigenic, grew faster and resulted in more metastases than did those occurring in mice inoculated with clones harboring other chromosomal abnormalities (Ghose *et al.*, 1990). This indicates that dup(1)(q11q32) provides a proliferative advantage. Furthermore, it is noteworthy that the minimally gained regions for both the BLs and the HeH ALLs in our study are included in the abovementioned segment. Hence, one or several genes mapping to the minimally gained region could be pathogenetically important.

The global gene expression analyses revealed five significantly up-regulated genes, namely *B4GALT3, DAP3, RGS16, TMEM183A* and *UCK2*. Two of them (*DAP3* and *UCK2*) were among the most overexpressed genes in the single BL case with gain of 1q investigated with expression array. The *DAP3* has been reported to be highly expressed in invasive glioblastoma multiforme cells (Mariani *et al.*, 2001), although it is normally proapoptotic. The UCK2 protein has been correlated with sensitivity to anticancer drugs, more specifically certain inhibitors of RNA polymerases (Shimamoto *et al.*, 2002). However, the impact of the up regulation of these genes in dup(1q)-positive high hyperdiploid ALLs and BLs is currently unknown and remains to be further investigated.

Article II

In *Article II*, an intriguing and unexpected finding was the differences between the HeH and the t(12;21)-positive ALLs as regards the chromosome-wide methylation patterns, with the majority of the commonly gained chromosomes

in the former subtype being less methylated than their disomic counterparts (**Figure 7B**). Only one previous study has reported similar findings (Weber *et al.*, 2005), identifying an overall lowered methylation in gene-poor regions of the trisomic chromosomes 7 and 14 in a colon cancer cell line. However, in the present study the decreased methylation was equally distributed, independent of gene content, which agrees well with previous data showing no clear bias for decreased methylation only of gene-poor regions in the inactive and globally hypomethylated X chromosome in females (Hellman and Chess, 2007). The inactive X chromosome was long believed to be hypermethylated, but was, quite unexpectedly, by the MeDIP approach demonstrated to be globally hypomethylated (Weber *et al.*, 2005).

It has previously been demonstrated that pediatric HeH ALL has a separate expression signature compared with other pediatric ALL subtypes (Andersson *et al.*, 2005) and that the majority of the genes on the gained chromosomes are upregulated compared to other ALL subgroups with no tri- or tetrasomies (**Figure 3**). In *Article II* we demonstrate a global overexpression of genes on the additional chromosomes also when comparing disomic and trisomic/tetrasomic chromosomes within the HeH group. Yeast experimental systems that promote aneuploidy also demonstrate expression of the majority of genes on the extra chromosomes (Torres *et al.*, 2007). However, even if the transcript levels are increased, the amount of detected proteins often is not. Thus, gene overexpression on additional chromosomes does not always translate directly into high amounts of protein. Hence, the role of the decreased methylation that we detected on the gained chromosomes in HeH could be unrelated to gene expression. Instead, these methylation differences could theoretically be involved in the compartmentalization of the tri-

/tetrasomic chromosomes during interphase, similar to Barr body formation of the inactive X chromosome. Another possibility is that the lowered methylation could contribute to the abrogation of the reduced cellular fitness that has been demonstrated to be associated with aneuploidy (Weaver *et al.*, 2007; Williams *et al.*, 2008).

The detection of subtype-specific methylation hotspots indicates that t(12;21)-positive and HeH ALLs have different global methylation patterns. Further support for this was derived from the clustering analyses which successfully could subgroup the t(12;21)-positive and HeH cases (**Figure 7A**). Thus, genome-wide methylation analyses can subdivide morphologically identical, but genetically and clinically distinct, hematologic malignancies.

One large methylation hotspot on 19q common to both subtypes has previously been reported in ALL. Interestingly, this chromosome segment harbors several maternally imprinted genes (Taylor *et al.*, 2007), indicating that other methylation hotspots identified in *Article II* possibly also could be associated with imprinted genes. To test this hypothesis we used available mapping information (Luedi *et al.*, 2007) on all known, or putatively, imprinted human genes, ascertaining the number of imprinted genes per chromosome band and how often bands with a specific number of imprinted genes harbored methylation hotspots. As seen in **Figure 9**, there was a high correlation between the presence of imprinted genes and methylation hotspots, with the frequencies of methylation hotspots increasing with the number of imprinted genes. This strongly suggests that regions rich in imprinted genes are associated with methylation hotspots in t(12;21)-positive and HeH ALLs. A possible explanation for this could be that imprinted genes

are more susceptible to *de novo* methylation of the active allele and hence biallelic silencing. But this remains to be elucidated.

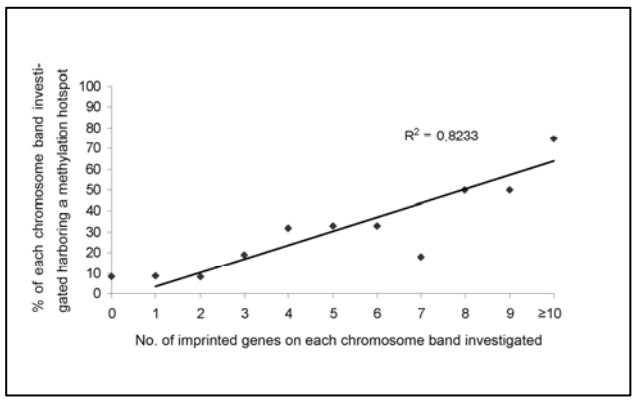


Figure 9. Association between methylation hotspots and imprinted gene density. Mapping information on all known, or putatively, imprinted human genes compared with the location of methylation hotspots. As seen, there is a high correlation between the presence of imprinted genes and hotspots, with the frequencies of hotspots increasing significantly with the number of imprinted genes.

Article III

In *Article III*, we investigated paired diagnostic/relapse samples using SNP array and mutation analysis of RTK-RAS genes. The array revealed higher frequencies of structural chromosome aberrations in relapse samples compared with diagnostic. This is well in line with previous findings, reporting increased numbers of structural abnormalities in HeH ALL at relapse (Secker-Walker *et al.*, 1989; Shikano *et al.*, 1990; Abshire *et al.*, 1992; Heerema *et al.*, 1992; Vora *et al.*, 1994; Chucrallah *et al.*, 1995; Mullighan *et al.*, 2008). We also detected more RTK-RAS mutations at relapse than at diagnosis. Whether this increase of genetic aberrations is caused by the chemotherapy administered to the patients or is a consequence of inherent clonal evolution over time remains unknown.

No structural aberration or mutated RTK-RAS gene was found in all cases. This indicates that relapse in HeH ALL most probably can not be related to one single genetic lesion, at least not detectable using current methods. Only four abnormalities, namely deletions involving *AK3*, *CDKN2A/B*, *IKZF1* or *PAX5*, were recurrent in the 11 investigated cases. Submicroscopic deletions in regions harboring these genes have previously been identified in HeH ALL at both diagnosis and relapse (Mullighan *et al.*, 2007; Mullighan *et al.*, 2008; Paulsson and Johansson, 2009). However, the prognostic impact of these aberrations is currently quite unclear (Den Boer *et al.*, 2009; Sulong *et al.*, 2009).

IKZF1 deletions have been shown to be associated with poor outcome in high-risk pediatric B-cell progenitor ALL (Mullighan *et al.*, 2009). In addition, specific SNPs in the gene have recently been associated with a increased risk of childhood ALL in two GWA studies (Papaemmanuil *et al.*, 2009; Treviño *et al.*, 2009). One of these *IKZF1* SNPs was demonstrated to generate a low *IKZF1* expression in Epstein-Barr virus-transformed lymphocytes (Papaemmanuil *et al.*, 2009), possibly mimicking a downregulation caused by a deletion. However, among the 11 patients investigated in the present series and the six paired samples previously reported by Mullighan *et al.* (2008), only 18% (n=3) harbored *IKZF1* deletions, with the losses being present both at diagnosis and relapse in one case, only at diagnosis in one and only at relapse in one. *IKZF1* deletions thus do not seem to play any significant prognostic role in the HeH genetic subgroup of pediatric ALL.

The simultaneous occurrence of +4, +10, and +17, called the "triple trisomies", is presently used to stratify low-risk childhood ALL by the COG

(Schultz *et al.*, 2007). It is hence noteworthy that we found these triple trisomies in the majority (64%) of the investigated paired diagnostic and relapse samples. The incidence of these triple trisomies in the close to 60 diagnostic HeH cases presently analyzed by high resolution arrays (Paulsson *et al.*, 2006; *Article I*; Mullighan *et al.*, 2007; Strefford *et al.*, 2007) is 50%, a frequency slightly lower than the one we observed in the paired samples. Furthermore, these trisomies were also observed in two of the eight previously reported paired diagnostic and relapse samples (Mullighan *et al.*, 2008; Yang *et al.*, 2008). It is thus quite clear that these "favorable" trisomies can be found in a substantial proportion of relapses in HeH, regardless of treatment protocols, thus questioning their favorable impact.

HeH ALL was genetically characterized only by the nonrandom tri- and tetrasomies and a handful of well-defined structural chromosome abnormalities before the advent of array based techniques (Paulsson and Johansson, 2009). However, the genetic patterns have, during recent years, been demonstrated to be more complex, consisting of several submicroscopic aberrations beside HeH and established cytogenetic abnormalities (Paulsson *et al.*, 2006; Mullighan *et al.*, 2007; Strefford *et al.*, 2007; Mullighan *et al.*, 2008). The complexity increases even further when paired diagnostic and relapse samples are compared. Our study (*Article III*) together with the one by Mullighan et al. (2008) demonstrates that the genetic relationship between diagnostic and relapse samples can be divided into different groups.

The most common group – the one displaying a complex genetic evolution pattern with some changes conserved and others lost or gained in the diagnostic and relapse samples, respectively – suggests the presence of an ancestral preleukemic clone from which both the diagnostic and relapse clones

are derived. Although it is not possible to directly characterize this clone genetically, its genotype can be backtracked by comparing diagnostic and relapsed samples from the same patient. When doing this it becomes clear that both structural chromosome changes and mutated genes are secondary to the HeH pattern, occurring later in the temporal order of the evolution of the neoplastic clone. Thus, HeH is the primary genetic abnormality in this ALL subtype.

Conclusions

Article I

- Proximal breakpoints of dup(1q)-positive BL and pediatric HeH ALL cluster close to the centromere in both disorders, suggesting this region to be breakprone.
- Distal breakpoints are more heterogeneous.
- The minimally gained region is much larger in the HeH ALLs than in the BLs.
- The satellite II domain is not hypomethylated in either disorder, ruling out that decondensation of pericentric heterochromatin, due to hypomethylation of satellite II sequences, is the underlying mechanism behind dup(1q) in BL and pediatric HeH ALL.
- Comparing HeH ALL dup(1q)-positive cases with HeH ALL without gain of 1q, five genes on 1q are significantly overexpressed and thus potentially pathogenetically important.

Article II

- Unsupervised cluster and principal component analyses of BAC array chromosome-wide methylome profiles can accurately subgroup the t(12;21) and HeH ALL subtypes, demonstrating that they are characterized by different methylomes.
- Analysis of the promoter-specific CpG island array demonstrates that several B-cell- and neoplasia-associated genes are hypermethylated and simultaneously underexpressed.
- Methylation hotspots are associated with chromosome bands predicted to harbor imprinted genes, suggesting that regions rich in imprinted genes are associated with aberrant methylation in t(12;21)-positive and HeH ALLs.
- The tri-/tetrasomic chromosomes in the HeH ALLs are less methylated than their disomic counterparts.

Article III

- Structural changes are significantly more common in relapse samples than in diagnostic ones.
- No single, recurrent aberration can be directly linked to relapse.
- The triple trisomies +4, +10, and +17, associated with low-risk ALL, are detected in two-thirds of cases that subsequently relapse.
- Genetic differences between diagnostic and relapse samples suggest the presence of an ancestral HeH clone in the vast majority of patients.
- Structural changes and mutations are secondary to the HeH pattern.

Concluding remarks

The major aim of this thesis was to analyze the genetic and epigenetic characteristics of pediatric HeH ALL using a spectrum of established, as well as novel techniques. Has this aim been fulfilled? The three original articles in this thesis have all used high-throughput genome-wide screening methods to address three different questions regarding HeH ALL: what are the genetic characteristics of the most common structural abnormality – dup(1q)? Are aberrant methylation patterns involved in leukemogenesis? How does diagnostic and relapse samples relate to each other genetically? In this fashion, a global genetic and epigenetic perspective has been obtained for these three relevant questions regarding HeH ALL pathogenesis.

Has any new significant insight regarding HeH leukemogenesis been gained? The fact that dup(1q) in HeH ALL probably is due to a breakprone region shared with other neoplasms harboring this abnormality indicates that this aberration is not specific for HeH but most likely contributes to neoplastic transformation in general. In the second study it was demonstrated that aberrant methylation resulting in downregulation of expression and that lowered methylation on the gained chromosomes could contribute to pathogenesis. The genetic differences between diagnostic and relapse samples in the third article suggested the presence of a preleukemic ancestral HeH clone giving rise to both diagnosis and relapse clones in the majority of cases. So, in essence, some important data contributing to the understanding of pediatric HeH ALL formation have been obtained. Of course, much work remains to be done in HeH ALL. For example, further investigation of the significance and frequency of different secondary genetic events contributing to overt disease needs to be undertaken in larger series using high-resolution

and genome-wide techniques. Validating a large proportion of the aberrantly methylated genes detected, as well as testing the functional outcome of their hypermethylation in experimental systems should definitely be performed. Investigating subclonality using SNP arrays, estimating allele frequencies of both diagnostic and relapsed samples, could give more clues to delineating the exact genetic patterns of ancestral clones. Last but definitely not least, the big question of the functional outcome of high hyperdiploidy as such still lingers. This is to a large extent unsolved and presently no experimental technique allows a direct investigation of these massive chromosomal imbalances. In resolving this lies the key to understanding the pathogenetics of pediatric HeH ALL.

SVENSK SAMMANFATTNING

(SUMMARY IN SWEDISH)

Hematologiska maligniteter är namnet på alla de kliniskt, morfologiskt och immunofenotypiskt heterogena cancersjukdomar som involverar benmärgen och blodet. De akuta leukemierna, vilka är de vanligaste formerna av hematologiska maligniteter hos barn och vuxna, delas in i de myeloiska och de lymfatiska, beroende på vilken blodcellslinje de har uppstått ifrån. Hos barn dominerar akut lymfatisk leukemi (ALL) och hos vuxna akut myeloisk leukemi. Båda karakteriseras av en ackumulering av maligna omogna vita blodkroppar i benmärg och perifert blod. Detta leder i sin tur till ökad infektionsbenägenhet och risk för blodbrist och blödning hos patienten.

Cancer är en sjukdom som karakteriseras av förvärvade genetiska förändringar i en cell, vilka ändrar dess egenskaper vad gäller tillväxt, utmognad och överlevnad. Vad gäller ALL hos barn, har förekomsten en karakteristisk topp mellan 2-5 år. Studier har visat att de första genetiska förändringarna ofta uppstår redan under fosterstadiet och att sedan ytterligare genetiska avvikelser ackumuleras i cellen innan leukemin kan bryta ut. Genetiska förändringar synbara på kromosomnivå detekteras i ungefär hälften av alla akuta leukemier och ger viktig klinisk information med betydelse för diagnostik och behandling. En vanlig strukturell genetisk avvikelse är en så kallad translokation, vilket innebär omflyttning av genetiskt material så att

cellens kontrollgener påverkas. Det är också vanligt att cellen får ett tillskott av kromosomer. Detta är dock inte lika välstuderat som translokationer.

Syftet med den här avhandlingen var att genetiskt karakterisera den vanligaste formen av barnleukemi: höghyperdiploid barn-ALL. Denna leukemityp karakteriseras främst av att cancercellerna innehåller ett stort antal extra kopior av vissa speciella kromosomer – så kallad hyperdiploidi. Idag vet vi väldigt lite vad detta kromosommönster har för betydelse och funktion i cancercellerna. Tidigare studier har visat att själva hyperdiploidin uppstår i fosterstadiet, genom en enda felaktig celldelning, och att ytterligare genetiska och epigenetiska (modifiering av genuttryck som inte beror på ändring av själva DNA sekvensen) förändringar sedan krävs för att en akut leukemi ska uppträda hos patienten. Höghyperdiploid barn-ALL fungerar således som en typisk modell för hur man idag tror naturalförloppet hos en barnleukemi är. Målet med de tre studierna som ingår i avhandlingen har varit att identifiera och karakterisera sekundära avvikelser till hyperdiploidin, avvikelser vilka kan ha betydelse för leukemiuppkomsten, men också för återinsjuknande.

I den första studien undersöktes tio stycken höghyperdiploida ALL och Burkitt lymfom, vilka bar på den vanligast förekommande strukturella cytogenetiska avvikelsen, nämligen överskott av kromosommaterial på den långa armen på kromosom 1, d.v.s. dup(1q). Detta gjordes med ”array CGH”, vilken är en teknik som detekterar väldigt små förluster eller tillkomster av genetiskt material. Vi kunde visa att en av de två brottspunkterna på duplikationerna var i samma region, nära kromosomens centrum, i samtliga fall. Det har spekulerats att dup(1q) kan bero på undermetylering av repetitiva DNA-sekvenser på denna kromosomarm. För att undersöka detta användes en metyleringskänslig Southern blot analys. Vi detekterade dock enbart normal

metylering av DNA-sekvenser på 1q, vilket indikerar att låg metyleringsgrad inte är mekanismen bakom dup(1q).

I studie II så undersökte vi metyleringsmönster över samtliga kromosomer och gener i 20 fall av höghyperdiploid och ALL med en translokation mellan kromosom 12 och 21, d.v.s. t(12;21). Med en antikropp fångade vi metylerade DNA-sekvenser och hybridiserade sedan dessa till en arrayplattform. Detta möjliggör undersökning av vilka gener och regioner som är högmetylerade, vilket ofta innebär att gener inte uttrycks. Följaktligen hittade vi också hypermetylering av ett stort antal gener i båda typerna av leukemi av vilka majoriteten också var nedreglerade uttrycksmässigt. Metylering är således ytterligare ett sätt att påverka gener i leukemiceller. Intressant nog kunde vi också visa en lägre metyleringsgrad i de extra kromosomkopiorna i höghyperdiploid jämfört med t(12;21)-positiv ALL.

I studie III undersökte vi genetiska skillnader mellan diagnos och recidiv hos 11 fall av höghyperdiploid barn-ALL. Metoderna var mutationsanalys och ”SNP array”, vilken likt array CGH detekterar genetiska obalanser på en hög resolutionsnivå. I linje med vad förgående undersökningar visat så hittade vi mer strukturella avvikelser i recidiv- gentemot diagnosproven. Dessutom fann vi att +4,+10 och +17, vilka har hävdats vara associerade med god prognos, i två tredjedelar av de undersökta fallen. Den genetiska jämförelsen visade också att den vanligaste utvecklingsvägen var från en gemensam höghyperdiploid preleukemisk klon, till vilken för diagnos och recidiv skilda strukturella avvikelser och mutation tillkommer. Sammanfattningsvis så har undersökningarna i denna avhandling lett till ökade insikter om det komplexa mönster av genetiska och epigenetiska förändringar som är associerade med utvecklingen av höghyperdiploid ALL.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude and appreciation to the following:

Bertil Johansson, my supervisor. Honestly, I could not have wished for a better tutor and supervisor! Your support, knowledge and enthusiasm have been a vital driving force in all my work associated with this thesis. Moreover, your integrity, dedication and passion when it comes to the medical profession are an inspiration. You are a true gentleman!

All my co-authors for your contributions; may it be samples or technical expertise, my studies would not have been possible without you!

Thoas Fioretos, my co-supervisor. For your invaluable scientific input when evaluating my projects and our co-authored papers.

Felix Mitelman, the former head of the Department of the Clinical Genetics, for accepting me as PhD student and creating a world class scientific environment based on support, instead of fierce competition.

Ulf Kristoffersson, head of the Lund Genetics Clinic, for supporting the department-clinic interaction and my EuroGentest work opportunities.

Bryan D Young, for kindly allowing me to work at his lab in London.

Kajsa Paulsson, my great partner-in-crime when it comes to HeH ALL!

Anna Andersson, for working with me on her valuable expression data.

Mattias Höglund, for introducing me to the Department.

Anette Welin, for solving everything. And then I mean <u>everything</u>.

Past and present PhD students and post docs: ***Björn Nilsson, David Lindgren, Emely Möller, Fransesca Micci, Gisela Lundberg, Hammurabi Bartuma, Helena Ågerstam, Henrik Lilljebjörn, Karin Rennstam, Karolin Hansén Nord, Kristina Karrman, Linda Holmquist Mengelbier, Malin Nilsson, Marcus Järås, Nils Hansen, Srinivas Veerla*** and ***Ylva Stewénius***. A special thanks to *Ylva* and *Hammurabi* for sharing desk with me and understanding that colloquial discussions and coffees are essential for my well being!

The senior staff: ***Anna Collin, Catarina Lundin, Charlotte Jin, David Gisselsson Nord, Fredrik Mertens, Ioannis Panagopoulos, Maria Soller, Markus Heidenblad, Nils Mandahl, Nina Larsson, Petra Johnels, Samuel Gebre-Medhin, Tord Jonsson,*** and ***Yusheng Jin***. A special thanks to *Mia, David* and *Anna* for helping me characterizing genetic syndromes.

The technical staff: ***Andrea Horvat, Carin Lassen, Jenny Nilsson, Linda Magnusson***, ***Margareth Isaksson*** and ***Marianne Rissler***. ***Bodil Strömbeck*** and ***Eva Pålsson*** might be enjoying retired life but are warmly remembered at the department!

All the staff at the clinic. I value your hard work and nice company.

The Oncology/BMC C13 people: ***Åke, Anna***, ***Göran***, ***Jeanette***, ***Joachim***, ***Johan S***, ***Johan V C*** and ***Markus***. It's been a pleasure lending your equipment and playing floorball with you.

All my friends outside the lab. Thank you for always being there, in good as well as bad times!

My family: Thank you ***Thage, Anna-Greta*** and ***Simon Davidsson*** for your care and support. Also, thanks to all my relatives in Lapland for accepting me as a frequent guest in your houses and beautiful nature.

Thank you all; family, friends and colleagues!

This research was made possible by grants from the *Swedish Childhood Cancer Foundation*, *the Swedish Cancer Society* and *the Swedish Research Council*.

REFERENCES

Abshire TC, Buchanan GR, Jackson JF, Shuster JJ, Brock B, Head D, Behm F, Crist WM, Link M, Borowitz M and Pullen DJ (1992). Morphologic, immunologic and cytogenetic studies in children with acute lymphoblastic leukemia at diagnosis and relapse: a Pediatric Oncology Group study. *Leukemia* **6**: 357-362.

Andersson A, Olofsson T, Lindgren D, Nilsson B, Ritz C, Edén P, Lassen C, Råde J, Fontes M, Mörse H, Heldrup J, Behrendtz M, Mitelman F, Höglund M, Johansson B and Fioretos T (2005). Molecular signatures in childhood acute leukemia and their correlations to expression patterns in normal hematopoietic subpopulations. *Proc Natl Acad Sci U S A* **102**: 19069-19074.

Armstrong SA, Mabon ME, Silverman LB, Li A, Gribben JG, Fox EA, Sallan SE and Korsmeyer SJ (2004). FLT3 mutations in childhood acute lymphoblastic leukemia. *Blood* **103**: 3544-3546.

Bamford S, Dawson E, Forbes S, Clements J, Pettett R, Dogan A, Flanagan A, Teague J, Futreal PA, Stratton MR and Wooster R (2004). The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. *Br J Cancer* **91**: 355-358.

Belkov VM, Krynetski EY, Schuetz JD, Yanishevski Y, Masson E, Mathew S, Raimondi S, Pui CH, Relling MV and Evans WE (1999). Reduced folate carrier expression in acute lymphoblastic leukemia: a mechanism for ploidy but not lineage differences in methotrexate accumulation. *Blood* **93**: 1643-1650.

Benson KF and Horwitz M (2006). Familial leukemia. *Best Pract Res Clin Haematol* **19**: 269-279.

Bernt KM and Armstrong SA (2009). Leukemia stem cells and human acute lymphoblastic leukemia. *Semin Hematol* **46**: 33-38.

Bhatia S, Sather HN, Heerema NA, Trigg ME, Gaynon PS and Robison LL (2002). Racial and ethnic differences in survival of children with acute lymphoblastic leukemia. *Blood* **100**: 1957-1964.

Bird A (2007). Perceptions of epigenetics. *Nature* **447**: 396-398.

Bock C, Reither S, Mikeska T, Paulsen M, Walter J and Lengauer T (2005). BiQ Analyzer: visualization and quality control for DNA methylation data from bisulfite sequencing. *Bioinformatics* **21**: 4067-4068.

Boultwood J and Wainscoat JS (2007). Gene silencing by DNA methylation in haematological malignancies. *Br J Haematol* **138**: 3-11.

Castor A, Nilsson L, Åstrand-Grundström I, Buitenhuis M, Ramirez C, Anderson K, Strömbeck B, Garwicz S, Békássy AN, Schmiegelow K, Lausen B, Hokland P, Lehmann S, Juliusson G, Johansson B and Jacobsen SE (2005). Distinct patterns of hematopoietic stem cell involvement in acute lymphoblastic leukemia. *Nat Med* **11**: 630-637.

Chucrallah AE, Stass SA, Huh YO, Albitar M and Kantarjian HM (1995). Adult acute lymphoblastic leukemia at relapse. Cytogenetic, immunophenotypic, and molecular changes. *Cancer* **76**: 985-991.

Clark SS, McLaughlin J, Crist WM, Champlin R and Witte ON (1987). Unique forms of the *abl* tyrosine kinase distinguish Ph1-positive CML from Ph1-positive ALL. *Science* **235**: 85-88.

Cooke HJ and Hindley J (1979). Cloning of human satellite III DNA: different components are on different chromosomes. *Nucleic Acids Res* **6**: 3177-3197.

Den Boer ML, van Slegtenhorst M, De Menezes RX, Cheok MH, Buijs-Gladdines JGCAM, Peters STCM, Van Zutven LJCM, Beverloo HB, Van der Spek PJ, Escherich G, Horstmann MA, Janka-Schaub GE, Kamps WA, Evans WE and Pieters R (2009). A subtype of childhood acute lymphoblastic leukaemia with poor treatment outcome: a genome-wide classification study. *Lancet Oncol* **10**: 125-134.

Feinberg AP and Vogelstein B (1983). Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* **301**: 89-92.

Feinberg AP and Tycko B (2004). The history of cancer epigenetics. *Nat Rev Cancer* **4**: 143-153.

Fiskus W, Wang Y, Sreekumar A, Buckley KM, Shi H, Jillella A, Ustun C, Rao R, Fernandez P, Chen J, Balusu R, Koul S, Atadja P, Marquez VE and Bhalla KN (2009). Combined epigenetic therapy with the histone methyltransferase EZH2 inhibitor 3-deazaneplanocin A and the histone deacetylase inhibitor panobinostat against human AML cells. *Blood* **114**: 2733-2743.

Forestier E, Johansson B, Borgström G, Kerndrup G, Johansson J and Heim S (2000a). Cytogenetic findings in a population-based series of 787 childhood acute lymphoblastic leukemias from the Nordic countries. The NOPHO Leukemia Cytogenetic Study Group. *Eur J Haematol* **64**: 194-200.

Forestier E, Johansson B, Gustafsson G, Borgström G, Kerndrup G, Johannsson J and Heim S (2000b). Prognostic impact of karyotypic findings in childhood acute lymphoblastic leukaemia: a Nordic series comparing two treatment periods. *Br J Haematol* **110**: 147-153.

Forestier E and Schmiegelow K (2006). The incidence peaks of the childhood acute leukemias reflect specific cytogenetic aberrations. *J Pediatr Hematol Oncol* **28**: 486-495.

Forestier E, Andersen MK, Autio K, Blennow E, Borgström G, Golovleva I, Heim S, Heinonen K, Hovland R, Johannsson JH, Kerndrup G, Nordgren A, Rosenquist R, Swolin B and Johansson B (2007). Cytogenetic patterns in *ETV6/RUNX1*-positive pediatric B-cell precursor acute lymphoblastic leukemia: A Nordic series of 245 cases and review of the literature. *Genes Chromosomes Cancer* **46**: 440-450.

Forestier E, Heyman M, Andersen MK, Autio K, Blennow E, Borgström G, Golovleva I, Heim S, Heinonen K, Hovland R, Johannsson JH, Kerndrup G, Nordgren A, Rosenquist R, Swolin B and Johansson B (2008a). Outcome of *ETV6/RUNX1*-positive childhood acute lymphoblastic leukaemia in the NOPHO-ALL-1992 protocol: frequent late relapses but good overall survival. *Br J Haematol* **140**: 665-672.

Forestier E, Izraeli S, Beverloo B, Haas O, Pession A, Michalová K, Stark B, Harrison CJ, Teigler-Schlegel A and Johansson B (2008b). Cytogenetic features of acute lymphoblastic and myeloid leukemias in pediatric patients with Down syndrome: an iBFM-SG study. *Blood* **111**: 1575-1583.

Garcia-Manero G, Yang H, Kuang SQ, O'Brien S, Thomas D and Kantarjian H (2009). Epigenetics of acute lymphocytic leukemia. *Semin Hematol* **46**: 24-32.

Ghose T, Lee CLY, Fernandez LA, Lee SHS, Raman R and Colp P (1990). Role of 1q trisomy in tumorigenicity, growth, and metastasis of human leukemic B-cell clones in nude mice. *Cancer Res* **50**: 3737-3742.

Golub TR, Barker GF, Bohlander SK, Hiebert SW, Ward DC, Bray-Ward P, Morgan E, Raimondi SC, Rowley JD and Gilliland DG (1995). Fusion of the *TEL* gene on 12p13 to the *AML1* gene on 21q22 in acute lymphoblastic leukemia. *Proc Natl Acad Sci U S A* **92**: 4917-4921.

Greaves M (2005). In utero origins of childhood leukaemia. *Early Hum Dev* **81**: 123-129.

Groupe Francais de Cytogénétique Hématologique (1993). Collaborative study of karyotypes in childhood acute lymphoblastic leukemias. *Leukemia* **7**: 10-19.

Hanahan D and Weinberg RA (2000). The hallmarks of cancer. *Cell* **100**: 57-70.

Harbott J, Viehmann S, Borkhardt A, Henze G and Lampert F (1997). Incidence of TEL/AML1 fusion gene analyzed consecutively in children with acute lymphoblastic leukemia in relapse. *Blood* **90**: 4933-4937.

Harrison CJ and Foroni L (2002). Cytogenetics and molecular genetics of acute lymphoblastic leukemia. *Rev Clin Exp Hematol* **6**: 91-113.

Harrison CJ and Johansson B (2009). Acute Lymphoblastic Leukemia. In: S. Heim and F. Mitelman (eds). *Cancer Cytogenetics*. Wiley-Blackwell, New Jersey**:** 233-296.

Hasle H (2001). Pattern of malignant disorders in individuals with Down's syndrome. *Lancet Oncol* **2**: 429-436.

Heerema NA, Palmer CG, Weetman R and Bertolone S (1992). Cytogenetic analysis in relapsed childhood acute lymphoblastic leukemia. *Leukemia* **6**: 185-192.

Heerema NA, Harbott J, Galimberti S, Camitta BM, Gaynon PS, Janka-Schaub G, Kamps W, Basso G, Pui C-H, Schrappe M, Auclerc M-F, Carroll AJ, Conter V, Harrison CJ, Pullen J, Raimondi SC, Richards S, Riehm H, Sather HN, Shuster JJ, Silverman LB, Valsecchi MG and Arico M (2004). Secondary cytogenetic aberrations in childhood Philadelphia chromosome positive acute lymphoblastic leukemia are nonrandom and may be associated with outcome. *Leukemia* **18**: 693-702.

Heinrichs S and Look AT (2007). Identification of structural aberrations in cancer by SNP array analysis. *Genome Biol* **8**: 219.

Hellman A and Chess A (2007). Gene body-specific methylation on the active X chromosome. *Science* **315**: 1141-1143.

Herman JG, Civin CI, Issa J-PJ, Collector MI, Sharkis SJ and Baylin SB (1997). Distinct patterns of inactivation of *p15^{INK4B}* and *p16^{INK4A}* characterize the major types of hematological malignancies. *Cancer Res* **57**: 837-841.

Hjalgrim LL, Rostgaard K, Schmiegelow K, Söderhall S, Kolmannskog S, Vettenranta K, Kristinsson J, Clausen N, Melbye M, Hjalgrim H and Gustafsson G (2003a). Age- and sex-specific incidence of childhood leukemia by immunophenotype in the Nordic countries. *J Natl Cancer Inst* **95**: 1539-1544.

Hjalgrim LL, Westergaard T, Rostgaard K, Schmiegelow K, Melbye M, Hjalgrim H and Engels EA (2003b). Birth weight as a risk factor for childhood leukemia: a meta-analysis of 18 epidemiologic studies. *Am J Epidemiol* **158**: 724-735.

Hjalgrim LL, Rostgaard K, Hjalgrim H, Westergaard T, Thomassen H, Forestier E, Gustafsson G, Kristinsson J, Melbye M and Schmiegelow K (2004). Birth weight and risk for childhood leukemia in Denmark, Sweden, Norway, and Iceland. *J Natl Cancer Inst* **96**: 1549-1556.

Hong D, Gupta R, Ancliff P, Atzberger A, Brown J, Soneji S, Green J, Colman S, Piacibello W, Buckle V, Tsuzuki S, Greaves M and Enver T (2008). Initiating and cancer-propagating cells in *TEL-AML1*-associated childhood leukemia. *Science* **319**: 336-339.

Horvath JE, Bailey JA, Locke DP and Eichler EE (2001). Lessons from the human genome: transitions between euchromatin and heterochromatin. *Hum Mol Genet* **10**: 2215-2223.

Hosack DA, Dennis GJ, Sherman BT, Lane HC and Lempicki RA (2003). Identifying biological themes within lists of genes with EASE. *Genome Biol* **4**: R70.

Hrusák O and Porwit-MacDonald A (2002). Antigen expression patterns reflecting genotype of acute leukemias. *Leukemia* **16**: 1233-1258.

Hunger SP (1996). Chromosomal translocations involving the E2A gene in acute lymphoblastic leukemia: clinical features and molecular pathogenesis. *Blood* **87**: 1211-1224.

Issa J-PJ and Kantarjian HM (2009). Targeting DNA methylation. *Clin Cancer Res* **15**: 3938-3946.

Jandl JH (1996). Acute Lymphatic Leukemia. In: *Blood: Textbook of Hematology, Second Edition*. Little, Brown and Company, Boston**:** 961-989.

Jeanpierre M, Turleau C, Aurias A, Prieur M, Ledeist F, Fischer A and Viegas-Pequignot E (1993). An embryonic-like methylation pattern of classical satellite DNA is observed in ICF syndrome. *Hum Mol Genet* **2**: 731-735.

Ji Y, Eichler EE, Schwartz S and Nicholls RD (2000). Structure of chromosomal duplicons and their role in mediating human genomic disorders. *Genome Res* **10**: 597-610.

Johansson B, Mertens F and Mitelman F (2004). Clinical and biological importance of cytogenetic abnormalities in childhood and adult acute lymphoblastic leukemia. *Ann Med* **36**: 492-503.

Jönsson G, Staaf J, Olsson E, Heidenblad M, Vallon-Christersson J, Osoegawa K, de Jong P, Oredsson S, Ringner M, Höglund M and Borg Å (2007). High-resolution genomic profiles of breast cancer cell lines assessed by tiling BAC array comparative genomic hybridization. *Genes Chromosomes Cancer* **46** 543-558.

Kager L, Cheok M, Yang W, Zaza G, Cheng Q, Panetta JC, Pui C-H, Downing JR, Relling MV and Evans WE (2005). Folate pathway gene expression differs in subtypes of acute lymphoblastic leukemia and influences methotrexate pharmacodynamics. *J Clin Invest* **115**: 110-117.

Kager L, Lion T, Attarbaschi A, Koenig M, Strehl S, Haas OA, Dworzak MN, Schrappe M, Gadner H and Mann G (2007). Incidence and outcome of *TCF3-PBX1*-positive acute lymphoblastic leukemia in Austrian children. *Haematologica* **92**: 1561-1564.

Kaneko Y, Hayashi Y and Sakurai M (1981). Chromosomal findings and their correlation to prognosis in acute lymphocytic leukemia. *Cancer Genet Cytogenet* **4**: 227-235.

Kaspers GJ, Smets LA, Pieters R, Van Zantwijk CH, Van Wering ER and Veerman AJ (1995). Favorable prognosis of hyperdiploid common acute lymphoblastic leukemia may be explained by sensitivity to antimetabolites and other drugs: results of an in vitro study. *Blood* **85**: 751-756.

Kasprzyk A, Harrison CJ and Secker-Walker LM (1999). Investigation of clonal involvement of myeloid cells in Philadelphia-positive and high hyperdiploid acute lymphoblastic leukemia. *Leukemia* **13**: 2000-2006.

Klco JM, Kreisel FH, Zehnbauer BA, Kulkarni S, Hassan A and Frater JL (2008). The spectrum of adult B-lymphoid leukemias with BCR-ABL: molecular diagnostic, cytogenetic, and clinical laboratory perspectives. *Am J Hematol* **83**: 901-907.

Kuiper RP, Schoenmakers EFPM, van Reijmersdal SV, Hehir-Kwa JY, Geurts van Kessel A, van Leeuwen FN and Hoogerbrugge PM (2007). High-resolution genomic profiling of childhood ALL reveals novel recurrent genetic lesions affecting pathways involved in lymphocyte differentiation and cell cycle progression. *Leukemia* **21**: 1258-1266.

Lampert F (1967). Cellulärer DNS-Gehalt und Chromozomenzahl bei der akuten Leukämie im Kindesalter und ihre Bedeutung für Chemotherapie und Prognose. *Klin Wschr* **45**: 763-768.

Larramendy M, El-Rifai W, Saarinen U, Alitalo R, Luomahaara S and Knuutila S (1995). Myeloid lineage involvement in acute lymphoblastic leukemia: a morphology antibody chromosomes (MAC) study. *Exp Hematol* **23**: 1563-1567.

le Viseur C, Hotfilder M, Bomken S, Wilson K, Röttgers S, Schrauder A, Rosemann A, Irving J, Stam RW, Shultz LD, Harbott J, Jürgens H, Schrappe M, Pieters R and Vormoor J (2008). In childhood acute lymphoblastic leukemia, blasts at different stages of immunophenotypic maturation have stem cell properties. *Cancer Cell* **14**: 47-58.

Li L-C and Dahiya R (2002). MethPrimer: designing primers for methylation PCRs. *Bioinformatics* **18**: 1427-1431.

Li Y, Qiu L, Zou D, Zhao Y, Mi Y and Wang J (2009). Additional chromosomal abnormalities and their prognostic significance in adult Philadelphia-positive acute lymphoblastic leukemia: with or without imatinib in chemotherapy. *Ann Hematol* **88**: 1069-1077.

Lin M, Wei L-J, Sellers WR, Lieberfarb M, Wong WH and Li C (2004). dChipSNP: significance curve and clustering of SNP-array-based loss-of-heterozygosity data. *Bioinformatics* **20**: 1233-1240.

Lingjaerde OC, Baumbusch LO, Liestøl K, Glad IK and Børresen-Dale A-L (2005). CGH-Explorer: a program for analysis of array-CGH data. *Bioinformatics* **21**: 821-822.

Lubbert M, Mertelsmann R and Herrmann F (1997). Cytosine methylation changes during normal hematopoiesis and in acute myeloid leukemia. *Leukemia* **11 Suppl 1**: S12-18.

Luedi PP, Dietrich FS, Weidman JR, Bosko JM, Jirtle RL and Hartemink AJ (2007). Computational and experimental identification of novel human imprinted genes. *Genome Res* **17**: 1723-1730.

Maia AT, van der Velden VHJ, Harrison CJ, Szczepanski T, Williams MD, Griffiths MJ, van Dongen JJM and Greaves MF (2003). Prenatal origin of hyperdiploid acute lymphoblastic leukemia in identical twins. *Leukemia* **17**: 2202-2206.

Maia AT, Tussiwand R, Cazzaniga G, Rebulla P, Colman S, Biondi A and Greaves M (2004). Identification of preleukemic precursors of hyperdiploid acute lymphoblastic leukemia in cord blood. *Genes Chromosomes Cancer* **40**: 38-43.

Mariani L, Beaudry C, McDonough WS, Hoelzinger DB, Kaczmarek E, Ponce F, Coons SW, Giese A, Seiler RW and Berens ME (2001). Death-associated protein 3 (Dap-3) is overexpressed in invasive glioblastoma cells in vivo and in glioma cell lines with induced motility phenotype in vitro. *Clin Cancer Res* **7**: 2480-2489.

Martineau M, Clark R, Farrell DM, Hawkins JM, Moorman AV and Secker-Walker LM (1996). Isochromosomes in acute lymphoblastic leukaemia: i(21q) is a significant finding. *Genes Chromosomes Cancer* **17**: 21-30.

Mitelman F, Johansson B and Mertens F (2009). Mitelman Database of Chromosome Aberrations in Cancer. *http://cgap.nci.nih.gov/Chromosomes/Mitelman.*

Molteni CG, Te Kronnie G, Bicciato S, Villa T, Tartaglia M, Basso G, Biondi A and Cazzaniga G (2009). *PTPN11* mutations in childhood acute lymphoblastic leukemia occur as a secondary event associated with high hyperdiploidy. *Leukemia* (in press).

Moorman AV, Richards SM, Martineau M, Luk Cheung K, Robinson HM, Reza Jalali G, Broadfield ZJ, Harris RL, Taylor KE, Gibson BES, Hann IM, Hill FGH, Kinsey SE, Eden TOB, Mitchell CD and Harrison CJ (2003). Outcome

heterogeneity in childhood high-hyperdiploid acute lymphoblastic leukemia. *Blood* **102**: 2756-2762.

Moorman AV, Raimondi SC, Pui C-H, Baruchel A, Biondi A, Carroll AJ, Forestier E, Gaynon PS, Harbott J, Harms DO, Heerema N, Pieters R, Schrappe M, Silverman LB, Vilmer E and Harrison CJ (2005). No prognostic effect of additional chromosomal abnormalities in children with acute lymphoblastic leukemia and 11q23 abnormalities. *Leukemia* **19**: 557-563.

Morison IM, Ellis LM, Teague LR and Reeve AE (2002). Preferential loss of maternal 9p alleles in childhood acute lymphoblastic leukemia. *Blood* **99**: 375-377.

Mullighan CG, Goorha S, Radtke I, Miller CB, Coustan-Smith E, Dalton JD, Girtman K, Mathew S, Ma J, Pounds SB, Su X, Pui C-H, Relling MV, Evans WE, Shurtleff SA and Downing JR (2007). Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* **446**: 758-764.

Mullighan CG, Phillips LA, Su X, Ma J, Miller CB, Shurtleff SA and Downing JR (2008). Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. *Science* **322**: 1377-1380.

Mullighan CG, Su X, Zhang J, Radtke I, Phillips LA, Miller CB, Ma J, Liu W, Cheng C, Schulman BA, Harvey RC, Chen IM, Clifford RJ, Carroll WL, Reaman G, Bowman WP, Devidas M, Gerhard DS, Yang W, Relling MV, Shurtleff SA, Campana D, Borowitz MJ, Pui C-H, Smith M, Hunger SP, Willman CL and Downing JR (2009). Deletion of *IKZF1* and prognosis in acute lymphoblastic leukemia. *N Engl J Med* **360**: 470-480.

Neff T and Armstrong SA (2009). Chromatin maps, histone modifications and leukemia. *Leukemia* **23**: 1243-1251.

Onodera N, McCabe NR and Rubin CM (1992). Formation of a hyperdiploid karyotype in childhood acute lymphoblastic leukemia. *Blood* **80**: 203-208.

Panzer-Grümayer ER, Fasching K, Panzer S, Hettinger K, Schmitt K, Stöckler-Ipsiroglu S and Haas OA (2002). Nondisjunction of chromosomes leading to hyperdiploid childhood B-cell precursor acute lymphoblastic leukemia is an early event during leukemogenesis. *Blood* **100**: 347-349.

Papaemmanuil E, Hosking FJ, Vijayakrishnan J, Price A, Olver B, Sheridan E, Kinsey SE, Lightfoot T, Roman E, Irving JAE, Allan JM, Tomlinson IP, Taylor M, Greaves M and Houlston RS (2009). Loci on 7p12.2, 10q21.2 and 14q11.2

are associated with risk of childhood acute lymphoblastic leukemia. *Nat Genet* **41**: 1006-1010.

Paulsson K, Panagopoulos I, Knuutila S, Jee KJ, Garwicz S, Fioretos T, Mitelman F and Johansson B (2003). Formation of trisomies and their parental origin in hyperdiploid childhood acute lymphoblastic leukemia. *Blood* **102**: 3010-3015.

Paulsson K, Mörse H, Fioretos T, Behrendtz M, Strömbeck B and Johansson B (2005). Evidence for a single-step mechanism in the origin of hyperdiploid childhood acute lymphoblastic leukemia. *Genes Chromosomes Cancer* **44**: 113-122.

Paulsson K, Heidenblad M, Mörse H, Borg Å, Fioretos T and Johansson B (2006). Identification of cryptic aberrations and characterization of translocation breakpoints using array CGH in high hyperdiploid childhood acute lymphoblastic leukemia. *Leukemia* **20**: 2002-2007.

Paulsson K, Horvat A, Strömbeck B, Nilsson F, Heldrup J, Behrendtz M, Forestier E, Andersson A, Fioretos T and Johansson B (2007). Mutations of *FLT3*, *NRAS*, *KRAS*, and *PTPN11* are frequent and possibly mutually exclusive in high hyperdiploid childhood acute lymphoblastic leukemia. *Genes Chromosomes Cancer* **47**: 26-33.

Paulsson K, Cazier JB, Macdougall F, Stevens J, Stasevich I, Vrcelj N, Chaplin T, Lillington DM, Lister TA and Young BD (2008). Microdeletions are a general feature of adult and adolescent acute lymphoblastic leukemia: unexpected similarities with pediatric disease. *Proc Natl Acad Sci U S A* **105**: 6708-6713.

Paulsson K, An Q, Moorman AV, Parker H, Molloy G, Davies T, Griffiths M, Ross FM, Irving J, Harrison CJ, Young BD and Strefford JC (2009). Methylation of tumour suppressor gene promoters in the presence and absence of transcriptional silencing in high hyperdiploid acute lymphoblastic leukaemia. *Br J Haematol* **144**: 838-847.

Paulsson K and Johansson B (2009). High hyperdiploid childhood acute lymphoblastic leukemia. *Genes Chromosomes Cancer* **48**: 637-660.

Pinkel D and Albertson DG (2005). Array comparative genomic hybridization and its applications in cancer. *Nat Genet* **37 Suppl**: S11-17.

Pollock BH, DeBaun MR, Camitta BM, Shuster JJ, Ravindranath Y, Pullen DJ, Land VJ, Mahoney Jr DH, Lauer SJ and Murphy SB (2000). Racial differences in the survival of childhood B-precursor acute lymphoblastic leukemia: a Pediatric Oncology Group Study. *J Clin Oncol* **18**: 813-823.

Pui C-H, Raimondi SC and Williams DL (1988). Isochromosome 17q in childhood acute lymphoblastic leukemia: an adverse cytogenetic feature in association with hyperdiploidy? *Leukemia* **2**: 222-225.

Pui C-H, Raimondi SC, Dodge RK, Rivera GK, Fuchs LA, Abromowitch M, Look AT, Furman WL, Crist WM and Williams DL (1989). Prognostic importance of structural chromosomal abnormalities in children with hyperdiploid (>50 chromosomes) acute lymphoblastic leukemia. *Blood* **73**: 1963-1967.

Pui C-H, Carroll AJ, Raimondi SC, Schell MJ, Head DR, Shuster JJ, Crist WM, Borowitz MJ, Link MP, Behm FG, Steuber CP and Land VJ (1992). Isochromosomes in childhood acute lymphoblastic leukemia: a collaborative study of 83 cases. *Blood* **79**: 2384-2391.

Pui C-H, Raimondi SC, Hancock ML, Rivera GK, Ribeiro RC, Mahmoud HH, Sandlund JT, Crist WM and Behm FG (1994). Immunologic, cytogenetic, and clinical characterization of childhood acute lymphoblastic leukemia with the t(1;19) (q23; p13) or its derivative. *J Clin Oncol* **12**: 2601-2606.

Pui C-H, Boyett JM, Hancock ML, Pratt CB, Meyer WH and Crist WM (1995). Outcome of treatment for childhood cancer in black as compared with white children. The St Jude Children's Research Hospital experience, 1962 through 1992. *JAMA* **273**: 633-637.

Pui C-H, Relling MV and Downing JR (2004). Acute lymphoblastic leukemia. *N Engl J Med* **350**: 1535-1548.

Pui C-H and Evans WE (2006). Treatment of acute lymphoblastic leukemia. *N Engl J Med* **354**: 166-178.

Pui C-H, Robison LL and Look AT (2008). Acute lymphoblastic leukaemia. *Lancet* **371**: 1030-1043.

Quijano CA, Moore D, 2nd, Arthur D, Feusner J, Winter SS and Pallavicini MG (1997). Cytogenetically aberrant cells are present in the $CD34^{+}CD33^{-}38^{-}19^{-}$ marrow compartment in children with acute lymphoblastic leukemia. *Leukemia* **11**: 1508-1515.

Rabbitts TH and Stocks MR (2003). Chromosomal translocation products engender new intracellular therapeutic technologies. *Nat Med* **9**: 383-386.

Raimondi SC, Pui C-H, Hancock ML, Behm FG, Filatov L and Rivera GK (1996). Heterogeneity of hyperdiploid (51-67) childhood acute lymphoblastic leukemia. *Leukemia* **10**: 213-224.

Romana SP, Le Coniat M and Berger R (1994). t(12;21): a new recurrent translocation in acute lymphoblastic leukemia. *Genes Chromosomes Cancer* **9**: 186-191.

Rowley JD (2008). Chromosomal translocations: revisited yet again. *Blood* **112**: 2183-2189.

Saal LH, Troein C, Vallon-Christersson J, Gruvberger S, Borg Å and Peterson C (2002). BioArray Software Environment (BASE): a platform for comprehensive management and analysis of microarray data. *Genome Biol* **3**: SOFTWARE0003.0001-0003.0006.

Sawyer JR, Tricot G, Mattox S, Jagannath S and Barlogie B (1998). Jumping translocations of chromosome 1q in multiple myeloma: evidence for a mechanism involving decondensation of pericentromeric heterochromatin. *Blood* **91**: 1732-1741.

Schmiegelow K and Hjalgrim H (2006). Is the risk of acute lymphoblastic leukemia reduced in siblings to children with the disease? A novel hypothesis explored by international collaboration. *Leukemia* **20**: 1206-1208.

Schmiegelow K, Al-Modhwahi I, Andersen MK, Behrendtz M, Forestier E, Hasle H, Heyman M, Kristinsson J, Nersting J, Nygaard R, Svendsen AL, Vettenranta K and Weinshilboum R (2009). Methotrexate/6-mercaptopurine maintenance therapy influences the risk of a second malignant neoplasm after childhood acute lymphoblastic leukemia: results from the NOPHO ALL-92 study. *Blood* **113**: 6077-6084.

Scholz C, Nimmrich I, Burger M, Becker E, Dorken B, Ludwig W-D and Maier S (2005). Distinction of acute lymphoblastic leukemia from acute myeloid leukemia through microarray-based DNA methylation analysis. *Ann Hematol* **84**: 236-244.

Schultz KR, Pullen DJ, Sather HN, Shuster JJ, Devidas M, Borowitz MJ, Carroll AJ, Heerema NA, Rubnitz JE, Loh ML, Raetz EA, Winick NJ, Hunger SP, Carroll WL, Gaynon PS and Camitta BM (2007). Risk- and response-based

classification of childhood B-precursor acute lymphoblastic leukemia: a combined analysis of prognostic markers from the Pediatric Oncology Group (POG) and Children's Cancer Group (CCG). *Blood* **109**: 926-935.

Secker-Walker LM, Alimena G, Bloomfield CD, Kaneko Y, Whang-Peng J, Arthur DC, de la Chapelle A, Reeves BR, Rowley JD, Lawler SD and Mitelman F (1989). Cytogenetic studies of 21 patients with acute lymphoblastic leukemia in relapse. *Cancer Genet Cytogenet* **40**: 163-169.

Shaffer LG, Slovak ML and Campbell LJ (2009). Neoplasia. In: *ISCN 2009*. Karger, Basel**:** 88-96.

Shaker S, Bernstein M, Momparler LF and Momparler RL (2003). Preclinical evaluation of antineoplastic activity of inhibitors of DNA methylation (5-aza-2'-deoxycytidine) and histone deacetylation (trichostatin A, depsipeptide) in combination against myeloid leukemic cells. *Leuk Res* **27**: 437-444.

Sharp AJ, Hansen S, Selzer RR, Cheng Z, Regan R, Hurst JA, Stewart H, Price SM, Blair E, Hennekam RC, Fitzpatrick CA, Segraves R, Richmond TA, Guiver C, Albertson DG, Pinkel D, Eis PS, Schwartz S, Knight SJL and Eichler EE (2006). Discovery of previously unidentified genomic disorders from the duplication architecture of the human genome. *Nat Genet* **38**: 1038-1042.

Shikano T, Ishikawa Y, Ohkawa M, Hatayama Y, Nakadate H, Hatae Y and Takeda T (1990). Karyotypic changes from initial diagnosis to relapse in childhood acute leukemia. *Leukemia* **4**: 419-422.

Shimamoto Y, Koizumi K, Okabe H, Kazuno H, Murakami Y, Nakagawa F, Matsuda A, Sasaki T and Fukushima M (2002). Sensitivity of human cancer cells to the new anticancer *ribo*-nucleoside TAS-106 is correlated with expression of uridine-cytidine kinase 2. *Jpn J Cancer Res* **93**: 825-833.

Smets LA, Slater R, van Wering ER, van der Does-van den Berg A, Hart AAM, Veerman AJP and Kamps WA (1995). DNA index and %S-phase cells determined in acute lymphoblastic leukemia of children: a report from studies ALL V, ALL VI, and ALL VII (1979-1991) of the Dutch Childhood Leukemia Study Group and The Netherlands Workgroup on Cancer Genetics and Cytogenetics. *Med Pediatr Oncol* **25**: 437-444.

Stam RW, den Boer ML, Schneider P, Meier M, Beverloo HB and Pieters R (2007). D-HPLC analysis of the entire *FLT3* gene in *MLL* rearranged and hyperdiploid acute lymphoblastic leukemia. *Haematologica* **92**: 1565-1568.

Strefford JC, Worley H, Barber K, Wright S, Stewart ARM, Robinson HM, Bettney G, van Delft FW, Atherton MG, Davies T, Griffiths M, Hing S, Ross FM, Talley P, Saha V, Moorman AV and Harrison CJ (2007). Genome complexity in acute lymphoblastic leukemia is revealed by array-based comparative genomic hybridization. *Oncogene* **26**: 4306-4318.

Sulong S, Moorman AV, Irving JAE, Strefford JC, Konn ZJ, Case MC, Minto L, Barber KE, Parker H, Wright SL, Stewart ARM, Bailey S, Bown NP, Hall AG and Harrison CJ (2009). A comprehensive analysis of the *CDKN2A* gene in childhood acute lymphoblastic leukemia reveals genomic deletion, copy number neutral loss of heterozygosity, and association with specific cytogenetic subgroups. *Blood* **113**: 100-107.

Surani MA, Barton SC and Norris ML (1984). Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nature* **308**: 548-550.

Suryanarayan K, Hunger SP, Kohler S, Carroll AJ, Crist W, Link MP and Cleary ML (1991). Consistent involvement of the *BCR* gene by 9;22 breakpoints in pediatric acute leukemias. *Blood* **77**: 324-330.

Suzuki MM and Bird A (2008). DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet* **9**: 465-476.

Synold TW, Relling MV, Boyett JM, Rivera GK, Sandlund JT, Mahmoud H, Crist WM, Pui C-H and Evans WE (1994). Blast cell methotrexate-polyglutamate accumulation in vivo differs by lineage, ploidy, and methotrexate dose in acute lymphoblastic leukemia. *J Clin Invest* **94**: 1996-2001.

Taketani T, Taki T, Sugita K, Furuichi Y, Ishii E, Hanada R, Tsuchida M, Sugita K, Ida K and Hayashi Y (2004). *FLT3* mutations in the activation loop of tyrosine kinase domain are frequently found in infant ALL with *MLL* rearrangements and pediatric ALL with hyperdiploidy. *Blood* **103**: 1085-1088.

Tartaglia M, Martinelli S, Cazzaniga G, Cordeddu V, Iavarone I, Spinelli M, Palmi C, Carta C, Pession A, Arico M, Masera G, Basso G, Sorcini M, Gelb BD and Biondi A (2004). Genetic evidence for lineage-related and differentiation stage-related contribution of somatic *PTPN11* mutations to leukemogenesis in childhood acute leukemia. *Blood* **104**: 307-313.

Taub JW, Konrad MA, Ge Y, Naber JM, Scott JS, Matherly LH and Ravindranath Y (2002). High frequency of leukemic clones in newborn screening blood

samples of children with B-precursor acute lymphoblastic leukemia. *Blood* **99**: 2992-2996.

Taylor GM, Dearden S, Ravetto P, Ayres M, Watson P, Hussain A, Greaves M, Alexander F and Eden OB (2002). Genetic susceptibility to childhood common acute lymphoblastic leukaemia is associated with polymorphic peptide-binding pocket profiles in *HLA-DPB1*0201*. *Hum Mol Genet* **11**: 1585-1597.

Taylor KH, Pena-Hernandez KE, Davis JW, Arthur GL, Duff DJ, Shi H, Rahmatpanah FB, Sjahputera O and Caldwell CW (2007). Large-scale CpG methylation analysis identifies novel candidate genes and reveals methylation hotspots in acute lymphoblastic leukemia. *Cancer Res* **67**: 2617-2625.

Taylor M, Harrison C, Eden T, Birch J, Greaves M, Lightfoot T and Hussain A (2008). *HLA-DPB1* supertype-associated protection from childhood leukaemia: relationship to leukaemia karyotype and implications for prevention. *Cancer Immunol Immunother* **57**: 53-61.

Taylor GM, Hussain A, Verhage V, Thompson PD, Fergusson WD, Watkins G, Lightfoot T, Harrison CJ and Birch JM (2009). Strong association of the *HLA-DP6* supertype with childhood leukaemia is due to a single allele, *DPB1*0601*. *Leukemia* **23**: 863-869.

Thompson JR, Gerald PF, Willoughby MLN and Armstrong BK (2001). Maternal folate supplementation in pregnancy and protection against acute lymphoblastic leukaemia in childhood: a case-control study. *Lancet* **358**: 1935-1940.

Torres EM, Sokolsky T, Tucker CM, Chan LY, Boselli M, Dunham MJ and Amon A (2007). Effects of aneuploidy on cellular physiology and cell division in haploid yeast. *Science* **317**: 916-924.

Treviño LR, Yang W, French D, Hunger SP, Carroll WL, Devidas M, Willman C, Neale G, Downing J, Raimondi SC, Pui C-H, Evans WE and Relling MV (2009). Germline genomic variants associated with childhood acute lymphoblastic leukemia. *Nat Genet* **41**: 1001-1005.

Uckun FM, Sensel MG, Sather HN, Gaynon PS, Arthur DC, Lange BJ, Steinherz PG, Kraft P, Hutchinson R, Nachman JB, Reaman GH and Heerema NA (1998). Clinical significance of translocation t(1;19) in childhood acute lymphoblastic

leukemia in the context of contemporary therapies: a report from the Children's Cancer Group. *J Clin Oncol* **16**: 527-535.

Wang X, Thomas P, Xue J and Fenech M (2004). Folate deficiency induces aneuploidy in human lymphocytes in vitro-evidence using cytokinesis-blocked cells and probes specific for chromosomes 17 and 21. *Mutat Res* **551**: 167-180.

Weaver BAA, Silk AD, Montagna C, Verdier-Pinard P and Cleveland DW (2007). Aneuploidy acts both oncogenically and as a tumor suppressor. *Cancer Cell* **11**: 25-36.

Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, Lam WL and Schubeler D (2005). Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat Genet* **37**: 853-862.

Venkatraman ES and Olshen AB (2007). A faster circular binary segmentation algorithm for the analysis of array CGH data. *Bioinformatics* **23**: 657-663.

Whitehead VM, Shuster JJ, Vuchich MJ, Mahoney DH, Jr., Lauer SJ, Payment C, Koch PA, Cooley LD, Look AT, Pullen DJ and Camitta B (2005). Accumulation of methotrexate and methotrexate polyglutamates in lymphoblasts and treatment outcome in children with B-progenitor-cell acute lymphoblastic leukemia: a Pediatric Oncology Group study. *Leukemia* **19**: 533-536.

Wiemels JL, Smith RN, Taylor GM, Eden OB, Alexander FE and Greaves MF (2001). Methylenetetrahydrofolate reductase (MTHFR) polymorphisms and risk of molecularly defined subtypes of childhood acute leukemia. *Proc Natl Acad Sci U S A* **98**: 4004-4009.

Wiemels JL, Zhang Y, Chang J, Zheng S, Metayer C, Zhang L, Smith MT, Ma X, Selvin S, Buffler PA and Wiencke JK (2005). *RAS* mutation is associated with hyperdiploidy and parental characteristics in pediatric acute lymphoblastic leukemia. *Leukemia* **19**: 415-419.

Williams BR, Prabhu VR, Hunter KE, Glazier CM, Whittaker CA, Housman DE and Amon A (2008). Aneuploidy affects proliferation and spontaneous immortalization in mammalian cells. *Science* **322**: 703-709.

Wong N, Lam W-C, Lai PB-S, Pang E, Lau W-Y and Johnson PJ (2001). Hypomethylation of chromosome 1 heterochromatin DNA correlates with q-arm copy gain in human hepatocellular carcinoma. *Am J Pathol* **159**: 465-471.

Vora AJ, Potter AM, Anderson LM and Lilleyman JS (1994). Frequency and importance of change in blast cell karyotype in relapsing childhood lymphoblastic leukemia. *Pediatr Hematol Oncol* **11**: 379-386.

Yagi T, Hibi S, Tabata Y, Kuriyama K, Teramura T, Hashida T, Shimizu Y, Takimoto T, Todo S, Sawada T and Imashuku S (2000). Detection of clonotypic IGH and TCR rearrangements in the neonatal blood spots of infants and children with B-cell precursor acute lymphoblastic leukemia. *Blood* **96**: 264-268.

Yang JJ, Bhojwani D, Yang W, Cai X, Stocco G, Crews K, Wang J, Morrison D, Devidas M, Hunger SP, Willman CL, Raetz EA, Pui C-H, Evans WE, Relling MV and Carroll WL (2008). Genome-wide copy number profiling reveals molecular evolution from diagnosis to relapse in childhood acute lymphoblastic leukemia. *Blood* **112**: 4178-4183.

Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J and Speed TP (2002). Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res* **30**: e15.

Zhang L, Taub JW, Williamson M, Wong SC, Hukku B, Pullen J, Ravindranath Y and Matherly LH (1998). Reduced folate carrier gene expression in childhood acute lymphoblastic leukemia: relationship to immunophenotype and ploidy. *Clin Cancer Res* **4**: 2169-2177.

I

Human Molecular Genetics, 2007, Vol. 16, No. 18 **2215–2225**
doi:10.1093/hmg/ddm173
Advance Access published on July 5, 2007

Tiling resolution array comparative genomic hybridization, expression and methylation analyses of dup(1q) in Burkitt lymphomas and pediatric high hyperdiploid acute lymphoblastic leukemias reveal clustered near-centromeric breakpoints and overexpression of genes in 1q22-32.3

Josef Davidsson[1,*], Anna Andersson[1], Kajsa Paulsson[1], Markus Heidenblad[1], Margareth Isaksson[1], Åke Borg[2,4], Jesper Heldrup[3], Mikael Behrendtz[5], Ioannis Panagopoulos[1], Thoas Fioretos[1] and Bertil Johansson[1]

[1]Department of Clinical Genetics, [2]Department of Oncology, and [3]Department of Pediatrics, Lund University Hospital, Lund University, SE-221 85 Lund, Sweden, [4]Lund Strategic Research Center for Stem Cell Biology and Cell Therapy, Lund University, Sweden and [5]Department of Pediatrics, Linköping University Hospital, SE-581 85 Linköping, Sweden

Received May 25, 2007; Revised and Accepted July 2, 2007

Although gain of 1q occurs in 25% of Burkitt lymphomas (BLs) and 10% of pediatric high hyperdiploid acute lymphoblastic leukemias (ALLs), little is known about the origin, molecular genetic characteristics and functional outcome of dup(1q) in these disorders. Ten dup(1q)-positive BLs/ALLs were investigated by tiling resolution (32k) array CGH analysis, which revealed that the proximal breakpoints in all cases were near-centromeric, in eight of them clustering within a 1.4 Mb segment in 1q12-21.1. The 1q distal breakpoints were heterogeneous, being more distal in the ALLs than in the BLs. The minimally gained segments in the ALLs and BLs were 57.4 Mb [dup(1)(q22q32.3)] and 35 Mb [dup(1)(q12q25.2)], respectively. Satellite II DNA on 1q was not hypomethylated, as ascertained by Southern blot analyses of 15 BLs/ALLs with and without gain of 1q, indicating that aberrant methylation was not involved in the origin of dup(1q), as previously suggested for other neoplasms with 1q rearrangements. Global gene expression analyses revealed that five genes in the minimally 57.4 Mb gained region—*B4GALT3*, *DAP3*, *RGS16*, *TMEM183A* and *UCK2*—were significantly overexpressed in dup(1q)-positive ALLs compared with high hyperdiploid ALLs without dup(1q). The *DAP3* and *UCK2* genes were among the most overexpressed genes in the BL case with gain of 1q investigated. The DAP3 protein has been reported to be highly expressed in invasive glioblastoma multiforme cells, whereas expression of the UCK2 protein has been correlated with sensitivity to anticancer drugs. However, involvement of these genes in dup(1q)-positive ALLs and BLs has previously not been reported.

INTRODUCTION

Gain of 1q, through duplication, isochromosome formation or unbalanced translocations, is one of the most frequent acquired cytogenetic abnormalities in human neoplasia (1). In solid tumors, dup(1q) is particularly common in hepatoblastoma (2), melanoma (3), Wilms' tumor (4) and carcinomas of the breast (5), liver (6), lung (7) and pancreas (8). As regards

*To whom correspondence should be addressed. Tel: +46 46173398; Fax: +46 46131061; Email: josef.davidsson@med.lu.se

hematological malignancies, dup(1q) has been reported in all major diagnostic subgroups (1), although mainly in B-lineage acute lymphoblastic leukemia (ALL) (9), multiple myeloma (MM) (10) and B-cell non-Hodgkin lymphoma (NHL) (11). In these disorders, the incidence of dup(1q) varies significantly among the different cytogenetic subgroups, being quite strongly associated with t(11;14)(q13;q32) in MM, t(3;14)(q27;q32), t(8;14)(q24;q32) and t(14;18)(q32;q21) in NHL and with high hyperdiploidy (>50 chromosomes) in pediatric ALL (1).

Despite the high prevalence of 1q gain in neoplasia, next to nothing is known about its origin, molecular genetic characteristics and functional outcome. However, methylation studies of Wilms' tumors and carcinomas of the breast, liver and ovary have revealed an association between hypomethylation of the heterochromatic 1q region and the presence of dup(1q), and it has been suggested that this undermethylation is the cause of the 1q rearrangements (12–15). Whether this holds true in hematological malignancies has not been investigated. In MM, a number of studies, using chromosome banding, fluorescence *in situ* hybridization (FISH), comparative genomic hybridization (CGH), single nucleotide polymorphism (SNP) and expression analyses, have revealed that the proximal breakpoints (PBs) often are near-centromeric and that several genes on 1q are overexpressed (16–20). The breakpoints in NHL, mainly follicular and diffuse large B-cell lymphomas, have also been shown to localize close to the centromere, often involving the satellite II (sat II) domain (17,21,22).

Surprisingly few studies have focused on dup(1q) in Burkitt lymphomas (BLs) and high hyperdiploid pediatric ALLs, which harbor 1q gains in approximately 25 and 10% of the cases, respectively (1,9,23–27). As in MM and NHL, the PBs in the relatively small number of dup(1q)-positive BLs analyzed have been shown to involve either the sat II domain or to be almost centromeric (17,21). To the best of our knowledge, no ALLs have been investigated in detail in this respect; however, a recent FISH mapping of two high hyperdiploid ALLs with 1q gain revealed near-centromeric PBs (28). In the present study, the PBs as well as the distal breakpoints (DBs) were analyzed in dup(1q)-positive high hyperdiploid ALLs and BLs, using tiling resolution (32k) array CGH. In addition, the methylation pattern of sat II was investigated in ALLs and BLs with and without dup(1q) in order to ascertain whether undermethylation of the heterochromatic region is associated with 1q rearrangements in these disorders. Finally, gene expression analyses, using 27k cDNA microarrays, were performed to elucidate the functional outcome of 1q gains in BLs and ALLs.

RESULTS

Cytogenetic characterization of 1q gains

The literature review revealed that the vast majority of the 1q PBs in dup(1q)-positive BLs and pediatric high hyperdiploid ALLs map cytogenetically close to the centromere, in 1q11-21, whereas the DBs are more heterogeneous. Although the duplications are quite large, often involving most of chromosome arm 1q, a minimally gained segment (1q23-25) could be delineated in both the disorders (Figs 1 and 2).

Array CGH findings

The array CGH analyses confirmed, refined and, in some cases, revised the aberrations identified by chromosome banding (Table 1). In addition, a total of 54 previously undetected abnormalities were revealed by the array CGH in the 10 cases investigated; many of the changes were submicroscopic, whereas a few larger ones most likely had escaped detection because of poor chromosome morphology. Of the 54 additional aberrations, 20 changes were potentially pathogenetic, whereas 25 involved regions known to harbor copy number polymorphisms (CNPs) and nine involved the *IGK@* and *IGH@* loci (Table 1). Excluding numerical aberrations, CNPs and the *IGK@* and *IGH@* rearrangements, 23 different genomic imbalances were found by array CGH (Table 2). All 54 imbalances in the 10 cases are summarized as a heat map in Figure 3.

The mapped 1q breakpoints and hence the sizes of the duplicated segments differed quite extensively between the chromosome banding and array CGH results (Table 1). Because of the often poor chromosome morphology in BLs/ALLs and the difficulty in delineating breakpoints cytogenetically, the findings from the array CGH analyses are most likely more accurate. As seen in Table 2, the 1q PBs in all 10 cases analyzed were near-centromeric, clustering within a 1.4 Mb segment at 1q12-21.1 (between chr 1:141.4 Mb and chr 1:142.8 Mb) in eight of the cases. The 1q DBs were more heterogeneous, generally being more distal in the ALLs than in the BLs. Although no clear-cut clustering was observed, three of the ALLs harbored DBs close to the telomere, at chr 1:245 Mb, whereas all four BLs had breakpoints between chr 1:176 Mb and chr 1:205 Mb (Table 2). The duplicated regions were hence larger in the ALLs (68.1–103.7 Mb) than in the BLs (35.4–63.6 Mb); the minimally gained segment in the ALLs was 57.4 Mb (chr 1:153.5–210.9 Mb), corresponding to dup(1)(q22q32.3), and 35 Mb (chr 1:141.8–176.8 Mb) in BLs, corresponding to dup(1)(q12q25.2) (Table 2; Figs 1 and 2).

Excluding CNPs and immunoglobulin rearrangements, additional cryptic, partial genomic imbalances were detected in two of the six ALLs, namely dup(4)(p16.2p16.3) and del(7)(q34) in case 1 and del(9)(q22.3q22.3) and del(12)(p12.1p13.2) in case 2 (Tables 1 and 2). In addition, the unbalanced 1;5 translocation in case 5 could be revised to der(5)t(1;5)(q22;q21.3) since the array CGH analysis revealed that the terminal deletion of 5q started in sub-band 5q21.3. Further FISH mapping of the 5q breakpoint, using the bacterial artificial chromosome probes RP11-541H04, RP11-505D04 and RP11-345I03 (BACPAC Resources, Oakland, CA, USA) revealed the presence of the first probe, located upstream of the *EFNA5* gene, weak signal of the second probe, covering *EFNA5* and loss of signal of the third probe, located downstream of the gene, strongly suggesting rearrangement of the *EFNA5* gene (data not shown). Despite several attempts to amplify a putative *EFNA5* fusion transcript using 5′-rapid amplification of cDNA ends (Clontech, Mountain View, CA, USA) no such chimera was obtained. The 7p breakpoint in the idic (7)(p11.2) in case 6, previously mapped in detail by locus-specific FISH (29), was confirmed by array CGH. Among the four BLs, cryptic imbalances were identified in two of them—dup(13)(q31.1q32.3) in case 8 and dup(X)

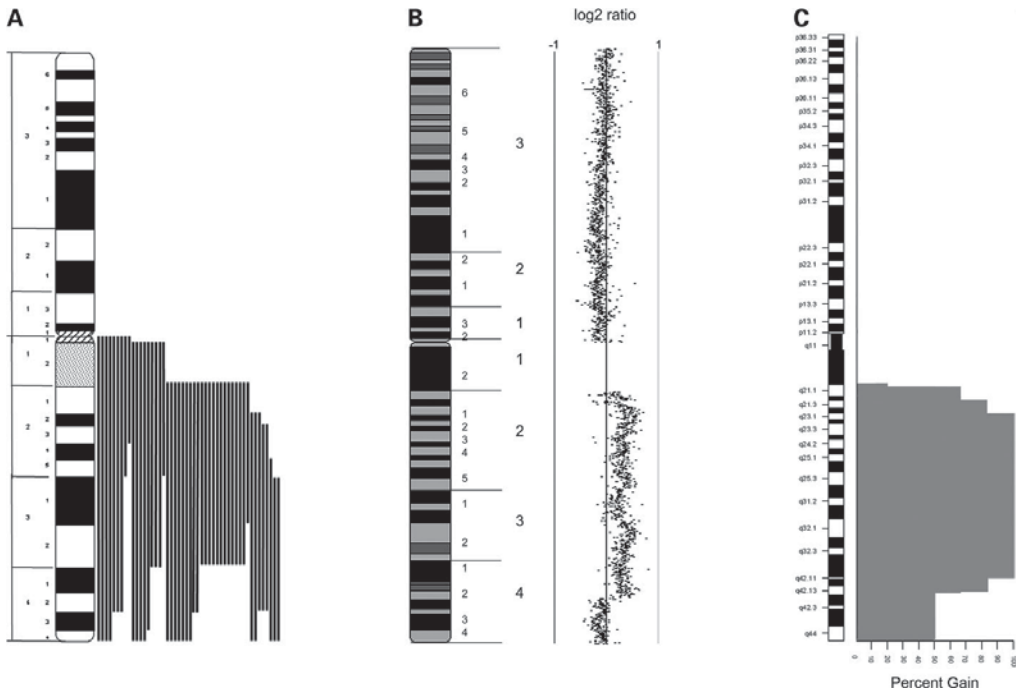


Figure 1. Gain of 1q in high hyperdiploid pediatric ALLs. (**A**) Cytogenetic imbalance map based on chromosome banding analyses of 48 cases reported in the literature (1) and unpublished cases from our department. (**B**) Representative example of array CGH 1q findings in ALL (case 1). (**C**) Frequency plot based on array CGH analysis of the 1q gains in cases 1–6.

(q28q28), dup(8)(q24.2q24.3), dup(14)(q32.3q32.3) and dup(18)(p11.2p11.3) in case 10. The presence of dup(8) and dup(14), involving regions in which *MYC* and some of the *IGH* genes map, in the latter case revealed that the cytogenetically balanced t(8;14) was unbalanced on the molecular genetic level. The add(17)(p13) in case 10 most likely represented a der(17)t(17;18)(p13.2;p11.2), considering the observed imbalances of these two chromosome arms. Unfortunately, lack of material precluded FISH confirmation.

Methylation status of sat II DNA

The methylation pattern of sat II DNA was investigated in a total of eight ALLs, at diagnosis and during remission; four of these ALLs harbored gain of 1q. Similarly, seven BLs, four of which with dup(1q), were also analyzed at diagnosis and during remission. Compared with a positive control for sat II hypomethylation and to normal peripheral blood DNA, the diagnostic and remission samples from all 15 cases studied showed a sat II methylation status similar to that observed in normal blood, and there were no clear-cut differences in intensities between the diagnostic and remission samples (Fig. 4). Hence, there was no evidence for sat II undermethylation in ALLs and BLs with or without 1q gains.

Expression patterns of genes in dup(1q)

The cDNA microarray analyses of 29 high hyperdiploid ALLs revealed that a large number of genes, mapping in the minimally gained 1q region identified herein (1q22-32.3), were highly expressed, irrespective of the genomic 1q status. In order to find genes in this chromosomal segment that were overexpressed in the dup(1q)-positive ALLs only, a *t*-test was performed, identifying six genes to be differentially expressed (*B4GALT3*, *CRB1*, *DAP3*, *RGS16*, *TMEM183A* and *UCK2*). All but one (*CRB1*) was highly expressed in the ALLs with 1q gain. Two of the genes were among the top 15 most highly expressed genes in the single dup(1q)-positive BL case analyzed, namely *DAP3* (rank 12) and *UCK2* (rank 3). The expression findings are summarized in Figure 5.

DISCUSSION

As expected, the array CGH analyses identified several imbalances in regions known to harbor CNPs (30) as well as

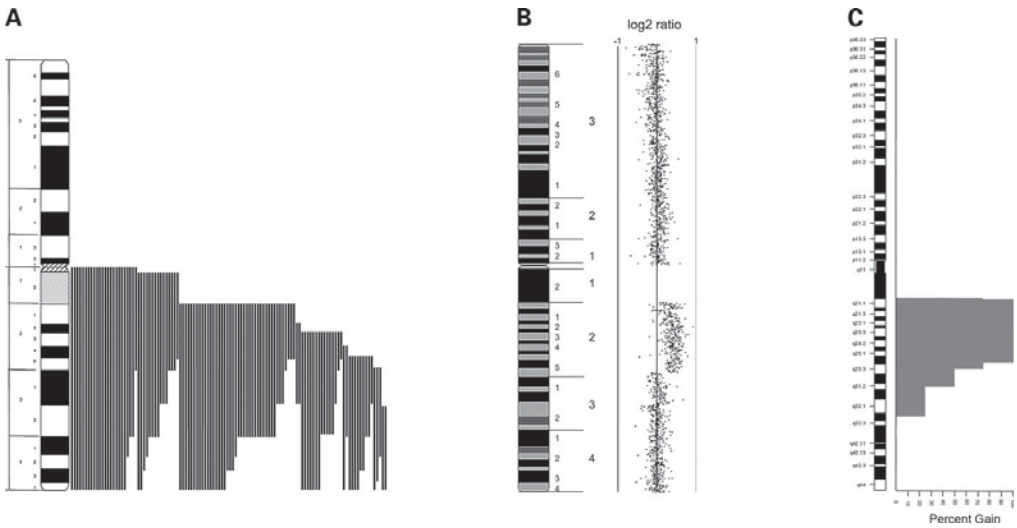


Figure 2. Gain of 1q in BLs. (**A**) Cytogenetic imbalance map based on G-banding analyses of 116 cases reported in the literature (1) and unpublished cases from our department. (**B**) Representative example of array CGH 1q findings in BL (case 8). (**C**) Frequency plot based on array CGH analysis of the 1q gains in cases 7–10.

Table 1. Clinical and cytogenetic features of the six high hyperdiploid ALLs and the four BLs with gain of 1q

Case	Diagnosis	Sex/age	Karyotypes based on chromosome banding (top row) and array CGH (lower row) analyses[a]
1[b,c]	ALL	M/7	50-54, XY, +4, +6, +10, +14, +17, +18, +20, +20, +mar 55, XY, +X, dup(1)(q21.1q42.1), +4, dup(4)(p16.2p16.3), +6, del(7)(q34), +10, **del(15)(q11.2q11.2)**, +17, +18, +18, **del(19)(p13.2p13.2)**, +21, +21
2[b]	ALL	M/11	57, XY, der(1)t(1;1)(p36;q21), +4, +6, +14, +17, +21, +21, inc 57, XY, +Y, dup(1)(q21.2q44), del(2)(p11.2p11.2), +4, +4, **del(4)(p11p11)**, +6, +8, del(9)(q22.3q22.3), +10, **del(11)(p14.1p14.1), del(11)(q22.1q22.1)**, del(12)(p12.1p13.2), +14, +17, +18, +21, +21, **del(22)(q11.2q11.2)**
3[b]	ALL	F/3	56, XX, +X, dup(1)(q?), +4, +6, +8, +10, +14, +14, +18, +21, +21 56, XX, +X, dup(1)(q21.1q32.3), del(2)(p11.2p11.2), +4, +6, **dup(7)(q22.1q22.1)**, +8, +10, +14, +14, +18, +21, +21
4[b]	ALL	M/3	52, XY, +X, dup(1)(q22q42), +6, +10, +11, +21, +21 52, XY, +X, dup(1)(q21.1q42.1), +6, **del(8)(p23.2p23.2)**, +10, +11, **dup(11)(p15.5p15.5)**, **dup(20)(q13.3q13.3)**, +21, +21
5[b]	ALL	M/2	61–63, XXY, +Y, −1, −2, −3, der(5)t(1;5)(p?;q?), −7, −13, +14, −15, −16, −19, −20, +21 61–63, XXY, +Y, −1, dup(1)(q22q44), −2, −3, del(5)(q21.3), −7, −13, +14, −15, **del(15)(q11.2q11.2)**, −16, **del(16)(p13.3)**, −19, −20, +21
6[b, c]	ALL	M/3	54–55, XY, +X, +4, +6, idic(7)(p11), +8, +10, +14, +17, +21, inc 54, XY, +X, dup(1)(q12q44), **dup(2)(p11.2p11.2)**, +4, +5, +6, del(7)(p11.2), dup(7)(q11.2q36.3), +8, +10, +14, **del(14)(q11.1q11.2)**, del(14)(q32.3q32.3), +17, +21
7	BL	M/7	46, XY, dup(1)(q25q42), t(8;14)(q24;q32)/46, XY, ?dup(1)(q21q11), t(8;14) 46, XY, dup(1)(q12q31.2), **dup(2)(p11.2p11.2)**
8	BL	F/39	46, XX, t(8;14)(q24;q32), der(13)t(1;13)(q21;q32)/46, idem, der(1)dup(1)(q?)add(1)(q42) 46, XX, dup(1)(q12q25.3), del(2)(p11.2p11.2), **del(4)(p16.3), del(5)(q31.3q31.3)**, **del(8)(q21.2q21.2)**, **del(10)(q26.3)**, **del(11)(p15.5)**, dup(13)(q31.1q32.3), del(13)(q32.3), del(14)(q32.3), **dup(15)(q11.1q11.2), del(22)(q11.2q11.2)**
9	BL	F/38	46, XX, dup(1)(q31q44), der(14)t(8;14)(q24;q32) 46, XX, dup(1)(q12q32.1), del(2)(p11.2p11.2), **dup(7)(q22.1q22.1)**, dup(8)(q24.2q24.3), del(14)(q32.3), **del(15)(q11.1q11.2)**
10	BL	M/45	46, XY, dup(1)(q25q31), t(8;14)(q24;q32), add(17)(p13) 46, dup(X)(q28q28), Y, dup(1)(q12q25.2), del(2)(p11.2p11.2), dup(8)(q24.2q24.3), dup(14)(q32.3q32.3), del(14)(q32.3), del(17)(p13.2), dup(18)(p11.2p11.3), **del(22)(q11.2q11.2)**

[a]Abnormalities in bold are suspected CNPs.
[b]Original karyotype previously published in (29,60–62).
[c]The gain of 1q in these cases was identified by SNP array analyses (unpublished data).

Table 2. Imbalances detected by array CGH in the 10 analyzed cases

Chromosome band[a]	Position (Mb)	Size (Mb)	Gain/Loss	Case(s)	Genes[b]	Clone[c]
1q12-q25.2	141.4–176.8	35.4	G	10		
1q12-q31.2	141.4–190.0	48.6	G	7		
1q12-q32.1	141.4–205.0	63.6	G	9		
1q12-q44	141.4–245.1	103.7	G	6		
1q12-q25.3	141.8–179.7	37.9	G	8		
1q21.1-q32.3	142.8–210.9	68.1	G	3		
1q21.1-q42.1	142.8–225.9	83.1	G	1,4		
1q21.2-q44	149.3–245.0	95.7	G	2		
1q22-q44	153.5–245.2	91.7	G	5		
2p11.2	87.2–88.6	1.4	G	6,7	*PLGLB1, PLGLB2, RGPD1, RGPD2, KRCC1, SMYD1, FABP1*	RP11-532O20
2p11.2	88.8–88.9	0.1	L	2, 3, 8–10	*IGK@*	RP11-15J7
4p16.3	0.0–1.0	1.0	L	8	*ZNF595, ZNF718, ZNF141, PIGG, PDE6B, ATP5I, MYL5, MFSD7, PCGF3, CPLX1, GAK, DGKQ, IDUA, FGFRL1*	RP11-467O4
4p16.2-p16.3	0.0–4.8	4.8	G	1		RP11-77M9
4p11	48.9–49.5	0.6	L	2	No known genes	RP11-445N8
5q21.3-q35.3	108.6–180.7	72.1	L	5		
5q31.3	140.0–140.4	0.4	L	8	*TMCO6, NDUFA2, IK, WDR55, DND1, HARS, HARSL, ZMAT2, PCDHA1-PCDHA13, PCDHAC1, PCDHAC2*	RP11-171H22
7p11.2-p22.3	0.0–56.2	56.2	L	6		
7q11.2-q36.3	61.8–158.6	96.8	G	6		
7q22.1	101.7–101.9	0.2	G	3,9	*CUTL1, PRKRIP1, ALKBH4*	RP13-771B7
7q34-q36.3	141.5–158.5	17.0	L	1		
8p23.2	3.7–5.2	1.5	L	4	*CSMD1*	RP11-626K14
8q21.2	86.6–87.0	0.4	L	8	*REXO1L1*	RP11-574H12
8q24.2-q24.3	128.8–146.2	17.4	G	9, 10		
9q22.3	95.6–98.2	2.6	L	2	*BARX1, PTPDC1, ZNF169, HIATL1, FBP2, FBP1, FANCC, PTCH1, HSD17B3, SLC35D2, ZNF367*	RP11-351M19
10q26.3	133.6–135.4	1.8	L	8	*PPP2R2D, BNIP3, DPYSL4, STK32C, LRRC27, PWWP2, INPPA, GPR123, KNDC1, UTF1, VENTX, ADAM8, TUBGCP2, ZNF511, DRD1IP, PRAP1, ECHS1, PAOX, MTG1, SPRN, CYP2E1, SYCE1, DUX4, FRG2*	RP11-692I8
11p15.5	0.0–2.7	2.7	G/L	4,8	*SCGB1C1, ODF3, BET1L, RIC8A, SIRT3, PSMD13, NALP6, ATHL1, IFITM1-IFITM3, IFITM5, B4GALNT4, PKP3, SIGIRR, TMEM16J, PTDSS2, RNH1, HRAS, LRRC56, RASSF7, IRF7, MUCDHL, SCT, DRD4, DEAF1, TMEM80, EPS8L2, TALDO1, PDDC1, BM88, LRDD, CEND1, LRDD, RPLP2, PNPLA2, EFCAB4A, CD151, POLR2L, TSPAN4, CHID1, AP2A2, MUC2, TOLLIP, BRSK2, HCCA2, DUSP8, FAM99A, CTSD, SYT8, TNNI2, LSP1, TNNT3, MRPL23, H19, IGF2, IGF2AS, INS, TH, ASCL2, TSPAN32, CD81, TSSC4, TRPM5, KCNQ1*	RP11-721I12
11p14.1	29.1–30.1	1.0	L	2	*KCNA4*	RP11-106M13
11q22.1	98.2–99.0	0.8	L	2	*CNTN5*	RP11-784P17
12p12.1-p13.2	10.4–22.8	12.4	L	2		
13q31.1-q32.3	85.2–100.3	15.1	G	8		
13q32.3-q34	100.3–114.1	13.8	L	8		
14q11.1-q11.2	18.1–19.4	1.3	L	6	*POTE14, OR1IH12, OR4Q3, OR4Q3, OR4M1, OR4N2*	RP11-558C21
14q32.3	102.5–105.3	2.8	G	10	*CDC42BPB, TNFAIP2, EIF5, MARK3, CKB, BAG5, KNS2, XRCC3, ZFYVE21, TDRD9, ADSSL1, SIVA, AKT1, PLD4, CDCA4, GPR132, JAG2, NUDT14, BRF1, BTBD6, PACS2, MTA1, CRIP2, CRIP1, TMEM121, IGHA1, IGHG1*	RP11-638H21
14q32.3	105.3–106.3	1.0	L	6,8–10	*IGHG2, IGHA1, IGHG1, IGH@*	RP11-761D24

Continued

Table 2. *Continued*

Chromosome band[a]	Position (Mb)	Size (Mb)	Gain/Loss	Case(s)	Genes[b]	Clone[c]
15q11.1-q11.2	18.3–22.9	4.6	G/L	1, 5, 8, 9		RP13-609E6
16p13.3	0.1–4.7	4.6	L	5		RP11-598P9
17p13.2-p13.3	0.0–4.9	4.9	L	10		RP11-229E7
18p11.2-p11.3	0.0–12.8	12.8	G	10		
19p13.2	8.7–8.8	0.1	L	1	*OR2Z1, ZNF558,*	RP11-680A23
20q13.3	61.2–62.4	1.2	G	4	*YTHDF1, BIRC7, ARFGAP1, COL20A1, CHRNA4, KCNQ2, EEF1A2, PTK6, SRMS, PRIC285, GMEB2, STMN3, RTEL1, TNFRSF6B, ARFRP1, ZGPAT, LIME1, BTBD4, ZBTB46, TPD52L2, DNAJC5, UCKL1, SAMD10, PRPF6, PRP17, SOX18, TCEA2, RGS19, OPRL1, NPBWR2, MYT1, PCMTD2*	RP11-1021O15
22q11.2	19.8–21.6	1.8	L	2, 8, 10	*GGT2, HIC2, MAPK1, PPM1F, TOP3B, IGLC1, VPREB1, SUHW2, PRAME, GGTL4*	RP11-757F24
Xq28	152.8–153.4	0.6	G	10	*LCAP, AVPR2, ARHGAP4, ARD1A, RENBP, HCFC1, IRAK1, MECP2, OPN1LW, OPN1MW2, OPN1MW, TKTL1, FLNA, EMD, RPL10, DNASE1L1, TAZ, ATP6AP1, GDI1, FAM50A, PLXNA3, LAGE3, UBLA4, FAM34*	RP11-333O6

[a]Chromosome bands in which CNPs have been described in two or more studies reported in the Database of Genomic Variants (http://projects.tcag.ca/variation/) are denoted in bold. CNPs were only considered for segments ≤3 Mb.
[b]Genes present in the region according to the USCS Genome Browser (http://genome.ucsc.edu/). Genes were only listed for segments ≤3 Mb.
[c]A representative clone is only given for segments ≤5 Mb.

deletions involving the *IGK@* and *IGH@* genes, the latter probably representing somatic immunoglobulin rearrangements clonotypic for the malignant lymphoid cells (31); these changes most likely did not contribute to the pathogenesis of the malignant disorders. On the other hand, submicroscopic, putative pathogenetic aberrations were found in two of four BLs and two of six pediatric ALLs. Similar high frequencies of additional changes, using array CGH, have previously been reported in high hyperdiploid ALLs (27,32) and in BLs (33,34). As regards BLs, none of the cryptic changes identified so far have been recurrent and their molecular genetic consequences remain to be elucidated, as do their clinical impact. Including the present series, array CGH findings in 15 high hyperdiploid ALLs, 12 of which analyzed with the 32k array set, have now been reported (27,32). Among these, a total of 14 additional genomic imbalances have been observed in seven cases. To date, only one change has been recurrent, namely a deletion of 12p including, among other genes, *ETV6*. Interestingly, very recently an SNP array analysis of 39 high hyperdiploid ALLs revealed *ETV6* deletions in two of the cases (35). Thus, cryptic hemizygous loss of this gene occurs in 5–10% of pediatric high hyperdiploid ALLs, suggesting that haploinsufficiency or loss of tumor suppressor function of *ETV6* may play a role in this common childhood leukemia.

The 1q PBs mapped close to the centromere in all BLs and ALLs investigated (Table 2 and Figs 1–3), a finding similar to what has been reported for other tumor types with 1q duplications (1). In the recent SNP array analysis of pediatric ALLs (35), approximately 30% of the high hyperdiploid cases were reported to harbor 1q gains. By using the publicly available dataset (http://ftp.stjude.org/pub/data/ALL-SNP1/) and the dchip software (36), the 1q PBs were extracted and shown to localize in the same region as the PBs in the present ALLs (data not shown). However, it should be noted that due to the repetitive nature of DNA sequences in this segment, the SNP density is lower than elsewhere in the genome, making accurate mapping difficult. This notwithstanding, the available data strongly indicate that the near-centromeric region of 1q is breakprone. The reason(s) for this is presently unknown, although it has been hypothesized that demethylation of the pericentric heterochromatin is associated with decondensation and subsequent instability of satellite sequences, resulting in 1q rearrangements (16), with several studies having reported undermethylation of sat II DNA in different neoplastic disorders (12–15). However, we found no evidence for hypomethylation of sat II in the present series of dup(1q)-positive BLs and ALLs (Fig. 4). Thus, we deem it highly unlikely that aberrant methylation of this region is the underlying mechanism behind dup(1q), at least not in BLs and ALLs. Instead, the genomic architecture as such, in particular segmental duplications, could be responsible. In fact, the presence of low-copy repeats has been implicated in the formation of both constitutional and neoplasia-associated chromosomal abnormalities (37–39). Interestingly, such duplications are very common immediately distal to highly repetitive satellite DNA close to the centromere, such as sat II (40), and recently (41–43), segmental duplications have been mapped to the 1q region in which the PBs occurred in our cases (Table 2). It is hence tempting

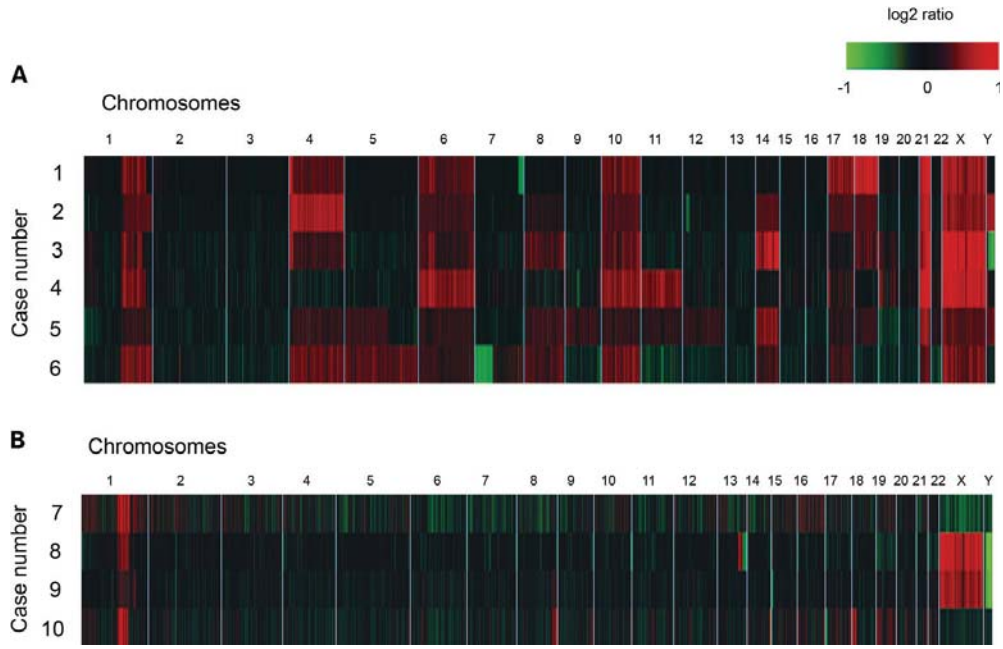


Figure 3. Heat map of genomic imbalances detected by array CGH analyses of six high hyperdiploid ALLs (cases 1–6) and four BLs (cases 7–10) with dup(1q).

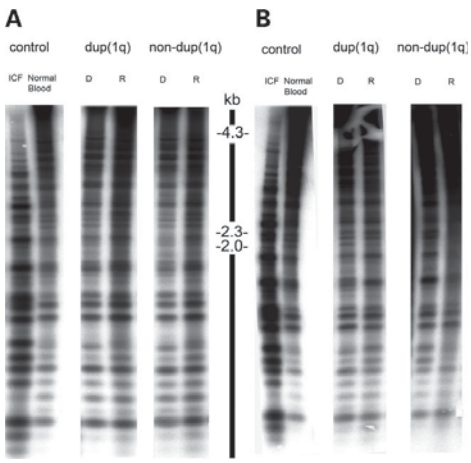


Figure 4. Representative Southern blot analyses of sat II DNA methylation with *Bst*BI, using DNA from ICF B-cells as a positive control for hypomethylation and DNA from normal peripheral blood as a reference and the sat II oligonucleotide probe. (**A**) Examples of high hyperdiploid ALLs, one with and one without dup(1q), at diagnosis (D) and at remission (R). (**B**) Similar examples of BLs.

to speculate that the dup(1q) in BLs and ALLs arises as a consequence of these rearrangement hotspots.

The 1q DBs were clearly more variable than the proximal ones, strongly indicating that there is not a common breakprone segment on distal 1q; in fact, none of the DBs (Table 2) are located in regions known to harbor such hotspots (42). Furthermore, the DBs also differed between the BLs and the ALLs, generally being more distal in the latter group (Table 2 and Figs 1–3). Thus, the minimally gained segment was larger in the ALLs (57.4 Mb) than in the BLs (35 Mb), corresponding to dup(1)(q22q32.3) and dup(1)(q12q25.2), respectively. Interestingly, in a previous array CGH analysis of high hyperdiploid ALLs (27) one case, without a cytogenetically identified dup(1q), harbored a submicroscopic 0.6 Mb gain of 1q22, close to the minimally gained segment identified herein. It has been reported that tumors arising in nude mice inoculated with a human leukemic B-cell clone carrying dup(1)(q11q32) were more tumorigenic, grew faster and resulted in more metastases than did those occurring in mice inoculated with clones harboring other chromosomal abnormalities, suggesting that gain of this 1q segment provides a proliferative advantage (44). It is noteworthy that the minimally gained regions in both the BLs and the ALLs are included in that segment. Hence, one or several genes mapping to the minimally gained region are potentially pathogenetically important.

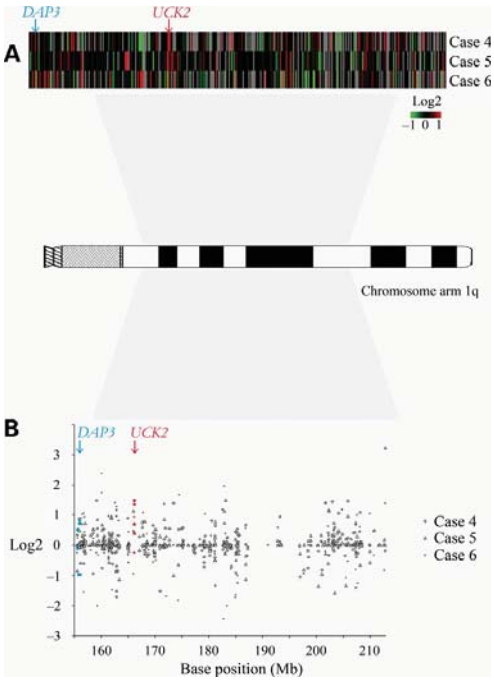


Figure 5. Expression profile of cDNA clones in the minimally gained 1q region in dup(1q)-positive high hyperdiploid ALLs. (**A**) Heat map of expressed cDNA clones in cases 4–6, with the location of the *DAP3* and *UCK2* genes, overexpressed also in a dup(1q)-positive BL, indicated. (**B**) Plot of cDNA clones in relation to Log2 ratios and genomic position on 1q. The *DAP3* and *UCK2* harboring clones are indicated in blue and red, respectively.

The microarray analyses revealed that six genes, namely *B4GALT3*, *CRB1*, *DAP3*, *RGS16*, *TMEM183A* and *UCK2*, were significantly differentially expressed in the dup(1q)-positive ALLs compared with the other high hyperdiploid cases. Surprisingly, one of these, the *CRB1* gene, was underexpressed. Thus, gain of chromosomal material does not necessarily lead to overexpression of all the duplicated genes, as previously reported also in other tumor types (45,46). As regards the five up-regulated genes, two of them, i.e. *DAP3* and *UCK2* (Fig. 5), were among the most overexpressed genes in the BL case with gain of 1q. Interestingly, the DAP3 (death associated protein 3) protein, which, if intact, is normally proapoptotic, has been reported to be highly expressed in invasive glioblastoma multiforme cells (47), whereas expression of the UCK2 protein has been correlated with sensitivity to a few anticancer drugs, i.e. certain inhibitors of RNA polymerases (48). However, the leukemogenic impact of these genes in dup(1q)-positive high hyperdiploid ALLs and BLs remains to be elucidated.

In conclusion, the salient findings in the present study of dup(1q)-positive BLs and pediatric high hyperdiploid ALLs, using tiling resolution array CGH, methylation and expression analyses, were that the PBs clustered close to the centromere in both these disorders, the DBs were more heterogeneous, the minimally gained region was much larger in the ALLs than in the BLs, the sat II domain was not hypomethylated, several additional changes were observed and that five genes on 1q were significantly overexpressed in the dup(1q)-positive cases compared with ALLs without gain of 1q.

MATERIALS AND METHODS

Patients

The array CGH study comprised six cases of B-cell precursor childhood high hyperdiploid ALL (five males, one female; age 2–11 years) and four cases of BL (two males, two females; age 7–45 years) (Table 1), all of which harbored 1q gains in the form of dup(1q) or unbalanced translocations involving 1q. Because of lack of material, FISH characterization of the 1q gains could only be performed in case 5. In the subsequent methylation investigations, diagnostic and remission DNA samples from cases 1–4 and 7–10 as well as from four additional high hyperdiploid ALLs and three BLs, without 1q gains, were analyzed. The gene expression patterns of cases 4–6 and 7 were extracted from a previously reported dataset (49).

Cytogenetic mapping of 1q gains

The Mitelman Database of Chromosome Aberrations in Cancer (1) was used to retrieve all cytogenetically characterized BLs and pediatric high hyperdiploid ALLs with 1q gains reported in the literature. In addition, unpublished cases from our department were included. In total, 116 BLs and 48 ALLs were identified. The PBs and DBs and the corresponding genomic imbalances were ascertained in all these cases.

Array CGH and gene expression analyses

DNA was extracted from bone marrow or peripheral blood cells obtained at the time of diagnosis and stored at −80°C. Male genomic DNA (Promega, Madison, WI, USA) was used as reference in all hybridizations. The 32k slides, containing 32 433 tiling bacterial artificial chromosome clones covering at least 98% of the human genome, were produced at the SWEGENE DNA microarray resource center at Lund University, Sweden. Labeling of DNA, slide preparation and hybridization were performed as described with minor modifications (50). Analyses of the microarray images were performed with the GenePix Pro 4.0 software (Axon Instruments, Foster City, CA, USA). For each spot, the median pixel intensity minus the median local background for both dyes was used to obtain the ratio of test gene copy number to reference gene copy number. Data normalization was performed for each array subgrid using lowess curve fitting with a smoothing factor of 0.33 (51). For all cases, the sex chromosomes and, for the high hyperdiploid ALLs, also the tri- and tetrasomic chromosomes were excluded when calculating the correction factor in the normalization. All analyses were performed in the Bioarray software environment (BASE) database

(52). To identify imbalances, the MATLAB toolbox CGH plotter (53), the SeeGH software (54) and the TM4 software suite (55) were applied using moving mean average over three clones and log2 limits of >0.2. Classification as gain or loss was based on identification as such by the CGH plotter and also by visual inspection of the log2 ratios. Ratios >0.5 in five adjacent clones were classified as aberrant, with ratios 0.5–1.0 interpreted as duplications/hemizygous deletions and >1.0 as amplifications/homozygous deletions.

The cDNA microarray analyses have previously been reported (49). In short, samples from 87 pediatric B-lineage ALLs, including 29 high hyperdiploid ALLs and two BLs, were hybridized to 27k microarray slides containing 25 648 cDNA clones (Swegene DNA Microarray Resource Center) representing 13 737 Unigene clusters and 11 592 Entrez gene entries, according to the Unigene build 195. RNA extraction, amplification, labeling, hybridization, scanning, post-hybridization washing and feature analysis were performed as described (49). The data analyses were all performed in BASE and in the CGH-Explorer software, and when investigating chromosome-specific expression patterns, genes were mean-centered and analyzed with a BASE implementation of the MATLAB toolbox CGH-plotter (53,56). Three of the six high hyperdiploid ALLs (cases 4–6) and one of the four BLs (case 7) with dup(1q) in the present study were included in the dataset. To identify differentially expressed genes in the dup(1q)-positive ALLs, a *t*-test comparing high hyperdiploid ALLs with and without gain of 1q was performed on the expression patterns of genes in the minimally gained 1q region. *P*-values <0.05 were considered significant.

Methylation analysis

Sat II methylation status was investigated by Southern blot analysis as described (12). Two micrograms of DNA were digested with 20 U of the CpG methyl-sensitive restriction enzyme *Bst*BI (New England BioLabs, Ipswich, MA, USA) for 18 h in standard conditions. The digested and fractionated DNA was blotted on a GeneScreen Plus hybridization transfer membrane (Perkin-Elmer, Waltham, MA, USA). The sat II probes used were an 18-mer single-stranded oligonucleotide, with the consensus sequence 5′-TCGAGTCCATTCGAT GAT-3′ (57), and a cloned Chr1-specific insert excised from the recombinant plasmid pUC1.77 (58). Hybridization was performed in Rapid-Hyb buffer (GE Healthcare Bio-Sciences Corporation, Piscataway, NJ, USA) at 42°C for the oligonucleotide probe and at 65°C for the cloned DNA probe. The oligonucleotide probe was radiolabeled with 5′-[γ^{32}P]-end dATP and the cloned Chr1-specific insert with α^{32}P-dCTP (GE Healthcare Bio-Sciences Corporation, Piscataway, NJ, USA). In each Southern blot analysis, normal peripheral blood and ICF B-lymphocyte (59) (Corriell Cell Repositories, Camden, NJ, USA) DNA was included as methylated and hypomethylated controls, respectively. Using a FLA 3000 phosphoimager (Fujifilm Corporation, Tokyo, Japan), the approximate amount of methylation was assessed by inspection, comparing the intensities of the hybridized fragments.

ACKNOWLEDGEMENTS

This work was supported by grants from the Swedish Cancer Society, the Swedish Children's Cancer Foundation, the Swedish Research Council and the Knut and Alice Wallenberg Foundation via the SWEGENE program.

Conflict of Interest statement. None declared.

REFERENCES

1. Mitelman, F., Johansson, B. and Mertens, F. (2007) Mitelman Database of Chromosome Aberrations in Cancer. http://cgap.nci.nih.gov/Chromosomes/Mitelman.
2. Tomlinson, G.E., Douglass, E.C., Pollock, B.H., Finegold, M.J. and Schneider, N.R. (2005) Cytogenetic evaluation of a large series of hepatoblastomas: numerical abnormalities with recurring aberrations involving 1q12-q21. *Genes Chromosomes Cancer*, **44**, 177–184.
3. Thompson, F.H., Emerson, J., Olson, S., Weinstein, R., Leavitt, S.A., Leong, S.P.L., Emerson, S., Trent, J.M., Nelson, M.A., Salmon, S.E. and Taetle, R. (1995) Cytogenetics of 158 patients with regional or disseminated melanoma. Subset analysis of near-diploid and simple karyotypes. *Cancer Genet. Cytogenet.*, **83**, 93–104.
4. Kullendorff, C.-M., Soller, M., Wiebe, T. and Mertens, F. (2003) Cytogenetic findings and clinical course in a consecutive series of Wilms tumors. *Cancer Genet. Cytogenet.*, **140**, 82–87.
5. Pandis, N., Jin, Y., Gorunova, L., Petersson, C., Bardi, G., Idvall, I., Johansson, B., Ingvar, C., Mandahl, N., Mitelman, F. and Heim, S. (1995) Chromosome analysis of 97 primary breast carcinomas: identification of eight karyotypic subgroups. *Genes Chromosomes Cancer*, **12**, 173–185.
6. Parada, L.A., Hallén, M., Tranberg, K.-G., Hägerstrand, I., Bondeson, L., Mitelman, F. and Johansson, B. (1998) Frequent rearrangements of chromosomes 1, 7 and 8 in primary liver cancer. *Genes Chromosomes Cancer*, **23**, 26–35.
7. Johansson, M., Berker Karaüzüm, S., Dietrich, C., Mandahl, N., Hambraeus, G., Johansson, L., Clausen, P.P., Mitelman, F. and Heim, S. (1994) Karyotypic abnormalities in adenocarcinomas of the lung. *Int. J. Oncol.*, **5**, 17–26.
8. Gorunova, L., Höglund, M., Andrén-Sandberg, Å., Dawiskiba, S., Jin, Y., Mitelman, F. and Johansson, B. (1998) Cytogenetic analysis of pancreatic carcinomas: intratumor heterogeneity and nonrandom pattern of chromosome aberrations. *Genes Chromosomes Cancer*, **23**, 81–99.
9. Johansson, B., Mertens, F. and Mitelman, F. (2004) Clinical and biological importance of cytogenetic abnormalities in childhood and adult acute lymphoblastic leukemia. *Ann. Med.*, **36**, 492–503.
10. Nilsson, T., Höglund, M., Lenhoff, S., Rylander, L., Turesson, I., Westin, J., Mitelman, F. and Johansson, B. (2003) A pooled analysis of karyotypic patterns, breakpoints and imbalances in 783 cytogenetically abnormal multiple myelomas reveals frequently involved chromosome segments as well as significant age- and sex-related differences. *Br. J. Haematol.*, **120**, 960–969.
11. Johansson, B., Mertens, F. and Mitelman, F. (1995) Cytogenetic evolution patterns in non-Hodgkin's lymphoma. *Blood*, **86**, 3905–3914.
12. Narayan, A., Ji, W., Zhang, X.Y., Marrogi, A., Graff, J.R., Baylin, S.B. and Ehrlich, M. (1998) Hypomethylation of pericentromeric DNA in breast adenocarcinomas. *Int. J. Cancer*, **77**, 833–838.
13. Wong, N., Lam, W.-C., Lai, P.B.-S., Pang, E., Lau, W.-Y. and Johnson, P.J. (2001) Hypomethylation of chromosome 1 heterochromatin DNA correlates with q-arm copy gain in human hepatocellular carcinoma. *Am. J. Pathol.*, **159**, 465–471.
14. Ehrlich, M., Hopkins, N.E., Jiang, G., Dome, J.S., Yu, M.C., Woods, C.B., Tomlinson, G.E., Chintagumpala, M., Champagne, M., Dillerg, L. *et al.* (2003) Satellite DNA hypomethylation in karyotyped Wilms tumors. *Cancer Genet. Cytogenet.*, **141**, 97–105.
15. Widschwendter, M., Jiang, G., Woods, C., Müller, H.M., Fiegl, H., Goebel, G., Marth, C., Müller-Holzner, E., Zeimet, A.G., Laird, P.W. and Ehrlich, M. (2004) DNA hypomethylation and ovarian cancer biology. *Cancer Res.*, **64**, 4472–4480.
16. Sawyer, J.R., Tricot, G., Mattox, S., Jagannath, S. and Barlogie, B. (1998) Jumping translocations of chromosome 1q in multiple myeloma: evidence for a mechanism involving decondensation of pericentromeric heterochromatin. *Blood*, **91**, 1732–1741.

17. Le Baccon, P., Leroux, D., Dascalescu, C., Duley, S., Marais, D., Esmenjaud, E., Sotto, J.J. and Callanan, M. (2001) Novel evidence of a role for chromosome 1 pericentric heterochromatin in the pathogenesis of B-cell lymphoma and multiple myeloma. *Genes Chromosomes Cancer*, **32**, 250–264.
18. Shaughnessy, J.D., Jr, Zhan, F., Burington, B.E., Huang, Y., Colla, S., Hanamura, I., Stewart, J.P., Kordsmeier, B., Randolph, C., Williams, D.R. *et al*. (2007) A validated gene expression model of high-risk multiple myeloma is defined by deregulated expression of genes mapping to chromosome 1. *Blood*, **109**, 2276–2284.
19. Walker, B.A., Leone, P.E., Jenner, M.W., Li, C., Gonzalez, D., Johnson, D.C., Ross, F.M., Davies, F.E. and Morgan, G.J. (2006) Integration of global SNP-based mapping and expression arrays reveals key regions, mechanisms, and genes important in the pathogenesis of multiple myeloma. *Blood*, **108**, 1733–1743.
20. Fabris, S., Ronchetti, D., Agnelli, L., Baldini, L., Morabito, F., Bicciato, S., Basso, D., Todoerti, K., Lombardi, L., Lambertenghi-Deliliers, G. and Neri, A. (2007) Transcriptional features of multiple myeloma patients with chromosome 1q gain. *Leukemia*, **21**, 1113–1116.
21. Busson-Le Coniat, M., Salomon-Nguyen, F., Dastugue, N., Maarek, O., Lafage-Pochitaloff, M., Mozziconacci, M.-J., Baranger, L., Brizard, F., Radford, I., Jeanpierre, M. *et al*. (1999) Fluorescence *in situ* hybridization analysis of chromosome 1 abnormalities in hematopoietic disorders: rearrangements of DNA satellite II and new recurrent translocations. *Leukemia*, **13**, 1975–1981.
22. Itoyama, T., Nanjungud, G., Chen, W., Dyomin, V.G., Teruya-Feldstein, J., Jhanwar, S.C., Zelenetz, A.D. and Chaganti, R.S.K. (2002) Molecular cytogenetic analysis of genomic instability at the 1q12-22 chromosomal site in B-cell non-Hodgkin lymphoma. *Genes Chromosomes Cancer*, **35**, 318–328.
23. Kornblau, S.M., Goodacre, A. and Cabanillas, F. (1991) Chromosomal abnormalities in adult non-endemic Burkitt's lymphoma and leukemia: 22 new reports and a review of 148 cases from the literature. *Hematol. Oncol.*, **9**, 63–78.
24. Raimondi, S.C., Pui, C.-H., Hancock, M.L., Behm, F.G., Filatov, L. and Rivera, G.K. (1996) Heterogeneity of hyperdiploid (51–67) childhood acute lymphoblastic leukemia. *Leukemia*, **10**, 213–224.
25. Moorman, A.V., Richards, S.M., Martineau, M., Cheung, K.L., Robinson, H.M., Jalali, G.R., Broadfield, Z.J., Harris, R.L., Taylor, K.E., Gibson, B.E.S. *et al*. (2003) Outcome heterogeneity in childhood high-hyperdiploid acute lymphoblastic leukemia. *Blood*, **102**, 2756–2762.
26. Lones, M.A., Sanger, W.G., Le Beau, M.M., Heerema, N.A., Sposto, R., Perkins, S.L., Buckley, J., Kadin, M.E., Kjeldsberg, C.R., Meadows, A. *et al*. (2004) Chromosome abnormalities may correlate with prognosis in Burkitt/Burkitt-like lymphomas of children and adolescents. A report from Children's Cancer Group Study CCG-E08. *J. Pediatr. Hematol. Oncol.*, **26**, 169–178.
27. Paulsson, K., Heidenblad, M., Mörse, H., Borg, Å., Fioretos, T. and Johansson, B. (2006) Identification of cryptic aberrations and characterization of translocation breakpoints using array CGH in high hyperdiploid childhood acute lymphoblastic leukemia. *Leukemia*, **20**, 2002–2007.
28. La Starza, R., Crescenzi, B., Pierini, V., Romoli, S., Gorello, P., Brandimarte, L., Matteucci, C., Kropp, M.G., Barbia, G., Martelli, M.F. and Mecucci, C. (2007) A common 93-kb duplicated DNA sequence at 1q21.2 in acute lymphoblastic leukemia and Burkitt lymphoma. *Cancer Genet. Cytogenet.*, **175**, 73–76.
29. Schaad, K., Strömbeck, B., Mandahl, N., Andersen, M.K., Heim, S., Mertens, F. and Johansson, B. (2006) FISH mapping of i(7q) in acute leukemias and myxoid liposarcoma reveals clustered breakpoints in 7p11.2: implications for formation and pathogenetic outcome of the idic(7)(p11.2). *Cytogenet. Genome Res.*, **114**, 126–130.
30. Redon, R., Ishikawa, S., Fitch, K.R., Feuk, L., Perry, G.H., Andrews, T.D., Fiegler, H., Shapero, M.H., Carson, A.R., Chen, W. *et al*. (2006) Global variation in copy number in the human genome. *Nature*, **444**, 444–454.
31. Wlodarska, I., Matthews, C., Veyt, E., Pospisilova, H., Catherwood, M.A., Poulsen, T.S., Vanhentenrijk, V., Ibbotson, R., Vandenberghe, P., Morris, T.C.M. and Alexander, H.D. (2007) Telomeric *IGH* losses detectable by fluorescence *in situ* hybridization in chronic lymphocytic leukemia reflect somatic V_H recombination events. *J. Mol. Diagn.*, **9**, 47–54.
32. Strefford, J.C., Worley, H., Barber, K., Wright, S., Stewart, A.R.M., Robinson, H.M., Bettney, G., van Delft, F.W., Atherton, M.G., Davies, T. *et al*. (2007) Genome complexity in acute lymphoblastic leukemia is revealed by array-based comparative genomic hybridization. *Oncogene*, **26**, 4306–4318.
33. Wessendorf, S., Schwaenen, C., Kohlhammer, H., Kienle, D., Wrobel, G., Barth, T.F.E., Nessling, M., Möller, P., Döhner, H., Lichter, P. and Bentz, M. (2003) Hidden gene amplifications in aggressive B-cell non-Hodgkin lymphomas detected by microarray-based comparative genomic hybridization. *Oncogene*, **22**, 1425–1429.
34. Hummel, M., Bentink, S., Berger, H., Klapper, W., Wessendorf, S., Barth, T.F.E., Bernd, H.-W., Cogliatti, S.B., Dierlamm, J., Feller, A.C. *et al*. (2006) A biologic definition of Burkitt's lymphoma from transcriptional and genomic profiling. *N. Engl. J. Med.*, **354**, 2419–2430.
35. Mullighan, C.G., Goorha, S., Radtke, I., Miller, C.B., Coustan-Smith, E., Dalton, J.D., Girtman, K., Mathew, S., Ma, J., Pounds, S.B. *et al*. (2007) Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature*, **446**, 758–764.
36. Schadt, E.E., Li, C., Ellis, B. and Wong, W.H. (2001) Feature extraction and normalization algorithms for high-density oligonucleotide gene expression array data. *J. Cell. Biochem. Suppl.*, **37**, 120–125.
37. Ji, Y., Eichler, E.E., Schwartz, S. and Nicholls, R.D. (2000) Structure of chromosomal duplicons and their role in mediating human genomic disorders. *Genome Res.*, **10**, 597–610.
38. Barbouti, A., Stankiewicz, P., Nusbaum, C., Cuomo, C., Cook, A., Höglund, M., Johansson, B., Hagemeijer, A., Park, S.-S., Mitelman, F. *et al*. (2004) The breakpoint region of the most common isochromosome, i(17q), in human neoplasia is characterized by a complex genomic architecture with large, palindromic, low-copy repeats. *Am. J. Hum. Genet.*, **74**, 1–10.
39. Fioretos, T. (2006) Mechanisms underlying neoplasia-associated genomic rearrangements. Lupski, J.R. and Stankiewicz, P. (eds), *The Genomic Basis of Disease*. Humana Press, Totowa, pp. 327–337.
40. Horvath, J.E., Bailey, J.A., Locke, D.P. and Eichler, E.E. (2001) Lessons from the human genome: transitions between euchromatin and heterochromatin. *Hum. Mol. Genet.*, **10**, 2215–2223.
41. Gregory, S.G., Barlow, K.F., McLay, K.E., Kaul, R., Swarbreck, D., Dunham, A., Scott, C.E., Howe, K.L., Woodfine, K., Spencer, C.C.A. *et al*. (2006) The DNA sequence and biological annotation of human chromosome 1. *Nature*, **441**, 315–321.
42. Sharp, A.J., Hansen, S., Selzer, R.R., Cheng, Z., Regan, R., Hurst, J.A., Stewart, H., Price, S.M., Blair, E., Hennekam, R.C. *et al*. (2006) Discovery of previously unidentified genomic disorders from the duplication architecture of the human genome. *Nat. Genet.*, **38**, 1038–1042.
43. Wong, K.K., deLeeuw, R.J., Dosanjh, N.S., Kimm, L.R., Cheng, Z., Horsman, D.E., MacAulay, C., Ng, R.T., Brown, C.J., Eichler, E.E. and Lam, W.L. (2007) A comprehensive analysis of common copy-number variations in the human genome. *Am. J. Hum. Genet.*, **80**, 91–104.
44. Ghose, T., Lee, C.L.Y., Fernandez, L.A., Lee, S.H.S., Raman, R. and Colp, P. (1990) Role of 1q trisomy in tumorigenicity, growth, and metastasis of human leukemic B-cell clones in nude mice. *Cancer Res.*, **50**, 3737–3742.
45. Pollack, J.R., Sørlie, T., Perou, C.M., Rees, C.A., Jeffrey, S.S., Lonning, P.E., Tibshirani, R., Botstein, D., Børresen-Dale, A-L. and Brown, P.O. (2002) Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. *Proc. Natl Acad. Sci. USA*, **99**, 12963–12968.
46. Storlazzi, C.T., Fioretos, T., Surace, C., Lonoce, A., Mastrorilli, A., Strömbeck, B., D'Addabbo, P., Iacovelli, F., Minervini, C., Aventin, A. *et al*. (2006) *MYC*-containing double minutes in hematologic malignancies: evidence in favor of the episome model and exclusion of *MYC* as the target gene. *Hum. Mol. Genet.*, **15**, 933–942.
47. Mariani, L., Beaudry, C., McDonough, W.S., Hoelzinger, D.B., Kaczmarek, E., Ponce, F., Coons, S.W., Giese, A., Seiler, R.W. and Berens, M.E. (2001) Death-associated protein 3 (Dap-3) is overexpressed in invasive glioblastoma cells *in vivo* and in glioma cell lines with induced motility phenotype *in vitro*. *Clin. Cancer Res.*, **7**, 2480–2489.
48. Shimamoto, Y., Koizumi, K., Okabe, H., Kazuno, H., Murakami, Y., Nakagawa, F., Matsuda, A., Sasaki, T. and Fukushima, M. (2002) Sensitivity of human cancer cells to the new anticancer ribo-nucleoside

TAS-106 is correlated with expression of uridine-cytidine kinase 2. *Jpn. J. Cancer Res.*, **93**, 825–833.
49. Andersson, A., Olofsson, T., Lindgren, D., Nilsson, B., Ritz, C., Edén, P., Lassen, C., Råde, J., Fontes, M., Mörse, H. *et al*. (2005) Molecular signatures in childhood acute leukemia and their correlations to expression patterns in normal hematopoietic subpopulations. *Proc. Natl Acad. Sci. USA*, **102**, 19069–19074.
50. Jönsson, G., Staaf, J., Olsson, E., Heidenblad, M., Vallon-Christersson, J., Osoegawa, K., de Jong, P., Oredsson, S., Ringnér, M., Höglund, M. and Borg, Å (2007) High-resolution genomic profiles of breast cancer cell lines assessed by tiling BAC array comparative genomic hybridization. *Genes Chromosomes Cancer*, **46**, 543–558.
51. Yang, Y.H., Dudoit, S., Luu, P., Lin, D.M., Peng, V., Ngai, J. and Speed, T.P. (2002) Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res.*, **30**, e15.
52. Saal, L.H., Troein, C., Vallon-Christersson, J., Gruvberger, S., Borg, Å. and Peterson, C. (2002) BioArray Software Environment (BASE): a platform for comprehensive management and analysis of microarray data. *Genome Biol.*, **3**, SOFTWARE0003.1-0003.6.
53. Autio, R., Hautaniemi, S., Kauraniemi, P., Yli-Harja, O., Astola, J., Wolf, M. and Kallioniemi, A. (2003) CGH-Plotter: MATLAB toolbox for CGH-data analysis. *Bioinformatics*, **19**, 1714–1715.
54. Chi, B., DeLeeuw, R.J., Coe, B.P., MacAulay, C. and Lam, W.L. (2004) SeeGH—a software tool for visualization of whole genome array comparative genomic hybridization data. *BMC Bioinformatics*, **5**, 13.
55. Saeed, A.I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., Braisted, J., Klapa, M., Currier, T., Thiagarajan, M. (2003) TM4: a free, open-source system for microarray data management and analysis. *Biotechniques*, **34**, 374–378.
56. Lingjærde, O.C., Baumbusch, L.O., Liestøl, K., Glad, I.K. and Børresen-Dale, L. (2005) CGH-Explorer: a program for analysis of array-CGH data. *Bioinformatics*, **21**, 821–822.
57. Tagarro, I., Fernández-Peralta, A.M. and González-Aguilera, J.J. (1994) Chromosomal localization of human satellites 2 and 3 by a FISH method using oligonucleotides as probes. *Hum. Genet.*, **93**, 383–388.
58. Cooke, H.J. and Hindley, J. (1979) Cloning of human satellite III DNA: different components are on different chromosomes. *Nucleic Acids Res.*, **6**, 3177–3197.
59. Jeanpierre, M., Turleau, C., Aurias, A., Prieur, M., Ledeist, F., Fischer, A. and Viegas-Pequignot, E. (1993) An embryonic-like methylation pattern of classical satellite DNA is observed in ICF syndrome. *Hum. Mol. Genet.*, **2**, 731–735.
60. Andreasson, P., Höglund, M., Békássy, A.N., Garwicz, S., Heldrup, J., Mitelman, F. and Johansson, B. (2000) Cytogenetic and FISH studies of a single center consecutive series of 152 childhood acute lymphoblastic leukemias. *Eur. J. Haematol.*, **65**, 40–51.
61. Paulsson, K., Panagopoulos, I., Knuutila, S., Jee, K.J., Garwicz, S., Fioretos, T., Mitelman, F. and Johansson, B. (2003) Formation of trisomies and their parental origin in hyperdiploid childhood acute lymphoblastic leukemia. *Blood*, **102**, 3010–3015.
62. Paulsson, K., Mörse, H., Fioretos, T., Behrendtz, M., Strömbeck, B. and Johansson, B. (2005) Evidence for a single-step mechanism in the origin of hyperdiploid childhood acute lymphoblastic leukemia. *Genes Chromosomes Cancer*, **44**, 113–122.

II

Human Molecular Genetics, 2009, Vol. 18, No. 21 **4054–4065**
doi:10.1093/hmg/ddp354
Advance Access published on August 13, 2009

The DNA methylome of pediatric acute lymphoblastic leukemia

Josef Davidsson[1,*], Henrik Lilljebjörn[1], Anna Andersson[1], Srinivas Veerla[1], Jesper Heldrup[2], Mikael Behrendtz[3], Thoas Fioretos[1] and Bertil Johansson[1]

[1]Department of Clinical Genetics and [2]Department of Pediatrics, Lund University Hospital, Lund University, SE-221 85 Lund, Sweden and [3]Department of Pediatrics, Linköping University Hospital, SE-581 85 Linköping, Sweden

Received June 23, 2009; Revised and Accepted July 26, 2009

Acute lymphoblastic leukemia (ALL) is the most common childhood malignancy, with high hyperdiploidy [51–67 chromosomes] and the t(12;21)(p13;q22) [*ETV6/RUNX1* fusion] representing the most frequent abnormalities. Although these arise *in utero*, there is long latency before overt ALL, showing that additional changes are needed. Gene dysregulation through hypermethylation may be such an event; however, this has not previously been investigated in a detailed fashion. We performed genome-wide methylation profiling using bacterial artificial chromosome arrays and promoter-specific analyses of high hyperdiploid and *ETV6/RUNX1*-positive ALLs. In addition, global gene expression analyses were performed to identify associated expression patterns. Unsupervised cluster and principal component analyses of the chromosome-wide methylome profiles could successfully subgroup the two genetic ALL types. Analysis of all currently known promoter-specific CpG islands demonstrated that several B-cell- and neoplasia-associated genes were hypermethylated and underexpressed, indicating that aberrant methylation plays a significant leukemogenic role. Interestingly, methylation hotspots were associated with chromosome bands predicted to harbor imprinted genes and the tri-/tetrasomic chromosomes in the high hyperdiploid ALLs were less methylated than their disomic counterparts. Decreased methylation of gained chromosomes is a previously unknown phenomenon that may have ramifications not only for the pathogenesis of high hyperdiploid ALL but also for other disorders with acquired or constitutional numerical chromosome anomalies.

INTRODUCTION

The two most common genetic subgroups of pediatric B-cell precursor acute lymphoblastic leukemia (ALL) are the ones represented by t(12;21)(p13;q22), leading to the *EVT6/RUNX1* fusion, or high hyperdiploidy [51–67 chromosomes], characterized by tri-/tetrasomies involving mainly chromosomes X, 4, 6, 8, 10, 14, 17, 18 and 21 (1). Although these genetic aberrations most likely are of utmost importance for leukemogenesis, there is ample evidence that they are not sufficient for leukemic transformation (2). The strongest support in favor of this has come from studies identifying *ETV6/RUNX1* or high hyperdiploidy in neonatal blood spots from children who later developed ALL. Thus, these changes occur prenatally but secondary events are needed for overt disease. Regarding *ETV6/RUNX1*, frequent additional changes include losses of the *EBF1*, *ETV6* and *PAX5* genes and gains of chromosome 21 and chromosome arm Xq, whereas common additional aberrations in high hyperdiploid ALL include duplication 1q, deletions of *CDKN2A* and *PAX5* and mutations of *FLT3*, *NRAS*, *KRAS* and *PTPN11* (3–8). Many of the secondary changes are genomic imbalances, aberrations that generally have been shown to affect gene dosage and lead to aberrant expression patterns (9). Gene dysregulation through epigenetic abnormalities may hence be an additional mechanism that also plays an important role. However, prior to the present study this has not been investigated in a detailed and genome-wide fashion in these common childhood ALL subtypes.

Hypermethylation of CpG islands situated within, or in close proximity of, promoters is an epigenetic modification frequently resulting in down-regulation or silencing of genes (10). To date, most methylation studies have been performed on individual genes in malignant solid tumors (11) (Supplementary Material, Table S1), and although several genes have been reported to be hypermethylated in B-lineage ALL (Supplementary Material,

*To whom correspondence should be addressed. Tel: +46 46173398; Fax: +46 46131061; Email: josef.davidsson@med.lu.se

Tables S2 and S3), such as *CDKN2B* and *FHIT*, nothing is known about the complete methylomes in this disorder. During recent years, several platforms assessing global CpG island methylation patterns have been developed (11–13). Until now, however, these methods have mainly been used on cell lines (13–17), and even though such analyses are important as a proof of principle, cell lines may have significantly altered methylomes due to *in vitro* culturing (15,18). As regards primary pediatric B-lineage ALL, most previous studies on aberrant methylation patterns have only focused on a few or even single genes (Supplementary Material, Table S3). In fact, we know of only two studies utilizing array-based methodologies to identify differential methylation of a larger number of genes in B-lineage ALL, one screening close to 60 genes (17) and one investigating ~2300 genes (16); however, these studies included mainly adult patients. Thus, next to nothing is known about the DNA methylome of childhood ALL.

Herein, we investigated primary patient samples from the two most common genetic subgroups of pediatric ALL, namely *ETV6/RUNX1* and high hyperdiploidy (Table 1). We analyzed large-scale global methylation profiles of each chromosome as well as direct gene silencing through hypermethylation by the use of a microarray platform comprising a total of 28 226 CpG islands in order to delineate the DNA methylome and to identify target genes for hypermethylation. In addition, gene expression data based on cDNA microarrays (19) were used to investigate correlations between hypermethylated target genes and their expression levels. Furthermore, tiling resolution array comparative genomic hybridization (array CGH) (20) was performed on each case to detect genomic imbalances that might affect gene expression and/or be associated with large-scale methylation changes.

RESULTS

Hypermethylated genes

A total of 8662 different genes with significant 5′ promoter CpG island methylation scores were identified by the CpG promoter array, constituting ~30% of all genes present on this platform. As seen in the flow chart in Figure 1, the high hyperdiploid ALLs harbored more hypermethylated genes than the *ETV6/RUNX1*-positive cases (7650 versus 3983), with 2971 being common to the two groups. The hypermethylated genes were distributed on all chromosomes, without any clear association with chromosomes frequently gained in *ETV6/RUNX1*-positive or high hyperdiploid ALLs (Supplementary Material, Table S4). The 30 genes displaying the highest methylation scores among all the 8662 hypermethylated genes are listed in Supplementary Material, Table S5 and the top 30 scored genes in the two genetic subtypes are given in Supplementary Material, Table S6 which includes six genes common to both cytogenetic subgroups. Supplementary Material, Table S7 lists only the top genes uniquely hypermethylated in each ALL subtype. As seen in these tables, the largest gene ontology (GO) group was transcription regulation.

Segmentation analysis of the CpG peak score data of the 8662 genes (see above and Fig. 1) identified 138 genes

Table 1. Clinical and cytogenetic features of the 20 analyzed pediatric ALLs with *ETV6/RUNX1* or high hyperdiploidy

Case no.	Sex/age (years)	Genetic subgroup	Genomic imbalances based on G-banding and array comparative genomic hybridization
1	M/5	*ETV6/RUNX1*	dup(X)(q21.31q28),dup(5)(q34q35.3),del(6)(q14.1q27),del(11)(q14.1q25)
2	M/9	*ETV6/RUNX1*	dup(12)(p13.2p13.33),dup(21)(q11.2q22.12)
3	M/5	*ETV6/RUNX1*	dup(X)(q25q28)
4	M/7	*ETV6/RUNX1*	del(6)(q14.1q22.1),del(6)(q24.3q27),dup(8)(q13.3q24.3),dup(12)(p13.2p13.33),del(13)(q13.1q31.1),dup(21)(q11.2q22.12)
5	F/4	*ETV6/RUNX1*	del(12)(p12.3p13.2),del(19)(q13.32q13.33)
6	M/8	*ETV6/RUNX1*	dup(10)(p11.1p15.3),+16,+21
7	F/3	*ETV6/RUNX1*	+X,del(18)(p11.23p11.32),dup(18)(p11.1p11.23),dup(18)(q11.1q23),+21
8	F/7	*ETV6/RUNX1*	−X,?del(1)(p36),-4,-5,del(6)(q?21),del(11)(p11),-12,del(21)(q22)
9	M/9	*ETV6/RUNX1*	dup(X)(q25q28),dup(12)(p13.2p13.33),dup(21)(q11.2q22.12)
10	M/14	*ETV6/RUNX1*	+5,del(9)(p21.3p21.3),del(9)(p21.3p22.1),dup(12)(p13.2p13.33),del(12)(p12.1p13.2),dup(21)(q11.2q22.12)
11	M/10	High hyperdiploid	+X,+6,dup(9)(p11.1p24.3),+10,+14,dup(14)(q11.2q11.2),+17,+18,+21,+21
12	M/3	High hyperdiploid	+X,dup(1)(q12q44),+4,+5,+6,del(7)(p11.2),dup(7)(q11.2q36.3),+8,+10,+14,+17,+21
13	F/3	High hyperdiploid	+X,+4,+6,+7,+10,+14,+17,+18,+21
14	F/1	High hyperdiploid	+X,+6,+8,+14,+17,+18,+21,+21
15	M/1	High hyperdiploid	+X,+4,+6,+14,+17,+18,+21,+21,+22
16	F/16	High hyperdiploid	+X,+4,+5,+6,+7,+8,del(9)(p21.3p21.3)x2,+10,+14,+14,+17,+18,+21,+21
17	F/4	High hyperdiploid	+X,+4,+6,+14,+17,+18,dup(18)(q22q23),+21,+21
18	M/2	High hyperdiploid	+X,+Y,dup(1)(q22q44),+4,dup(5)(p11p15.3),dup(5)(q11q21.2),+6,+8,+9,+10,+11,+12,+14,+14,+17,+18,+21,+21,+22
19	F/4	High hyperdiploid	+X,+4,+8,+10,+18,+21
20	F/6	High hyperdiploid	+X,+4,+6,+14,+17,+21

The original karyotypes of all cases and the array CGH data on cases 1–10, 12 and 18 have previously been reported (5,7,8).

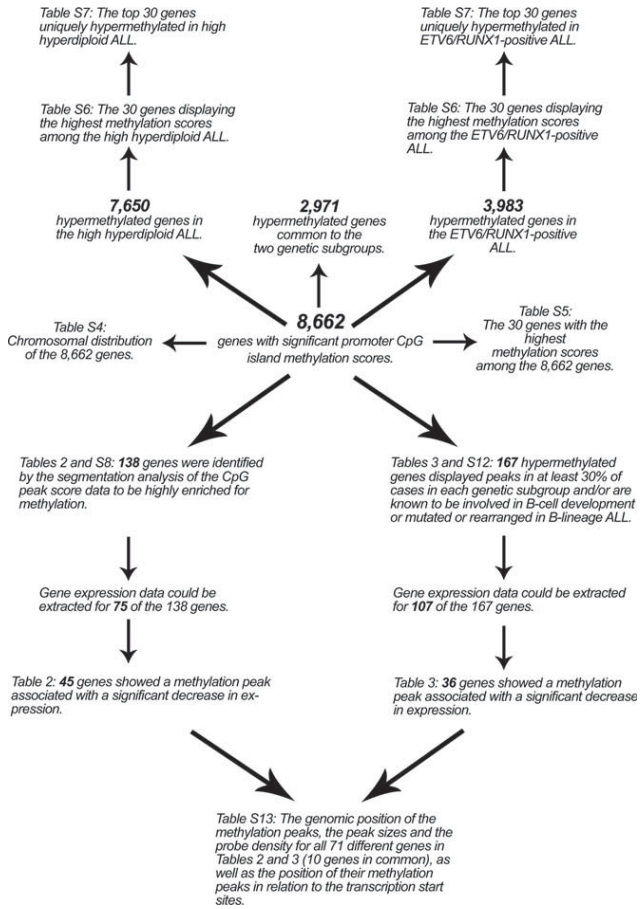


Figure 1. Flow chart showing how the 8662 genes with significant promoter CpG island methylation scores were grouped according to genetic ALL subtype, chromosomal distribution, peak scores and effect on gene expression.

that were highly enriched for methylation (Supplementary Material, Fig. S1). These genes could be subgrouped into seven GO groups, with the most common being cell signaling, transcription regulation and apoptosis. Gene expression data, corresponding to log2 ratios for $>50\%$ of cases in each genetic subgroup, could be extracted for 75 of the 138 genes. It was apparent that enrichment for methylation not necessarily resulted in decreased expression. In fact, 11 of the 75 genes displayed a higher mean expression level in cases with methylation peaks than in those without such peaks, whereas 19 did not differ substantially between methylated and non-methylated cases (Supplementary Material, Table S8). However, the majority (45 genes; 60%) showed a peak associated with a significant decrease in expression (Table 2).

A total of 117 genes displayed peaks in $\geq 30\%$ of cases in each genetic subgroup, and among the 236 different genes listed in Supplementary Material, Tables S1, S2 and S9–S11, 97 displayed CpG island peaks, 47 of which were also included among the 117 methylated genes identified earlier. Thus, a total of 167 different recurring genes and/or genes involved in B-cell development or mutated or rearranged in B-lineage ALL were found to be targets for hypermethylation. Only 18 of these were among the 138 genes found by the segmentation analysis (see above and Fig. 1). The 167 genes could be subgrouped into nine specific GO groups, with the most common being transcription regulation, apoptosis and cell signaling (Supplementary Material, Fig. S2). Gene expression data could be extracted for 107 of the 167 genes. Thirteen of the 107 genes displayed a higher mean expression

Table 2. The 45 genes shown by segmental analysis to be highly enriched for methylation and that displayed significantly[a] decreased expression in *ETV6/RUNX1*-positive and/or high hyperdiploid cases

Gene	Location[b]	Function[b]	Genetic subgroup[c]
BRDT	1p22.1	Transcription regulation	*ETV6/RUNX1*
ETV1	7p21.3	Transcription regulation	*ETV6/RUNX1*
HLF	17q22	Transcription regulation	High hyperdiploid
PARP14	3q21.1	Transcription regulation	High hyperdiploid
SOX30	5q33	Transcription regulation	*ETV6/RUNX1*, high hyperdiploid
SSX1	Xp11.22–11.23	Transcription regulation	*ETV6/RUNX1*
SSX4	Xp11.23	Transcription regulation	High hyperdiploid
TULP4	6q25–26	Transcription regulation	High hyperdiploid
CSF2RA	Xp22.32, Yp11.3	Cell signaling	*ETV6/RUNX1*, high hyperdiploid
DUSP10	1q41	Cell signaling	High hyperdiploid
HRH1	3p25	Cell signaling	High hyperdiploid
PAQR3	4q21.21	Cell signaling	High hyperdiploid
PRKD3	2p21	Cell signaling	High hyperdiploid
PPFIA4	1q32.1	Cell signaling	High hyperdiploid
PTPN3	9q31	Cell signaling	*ETV6/RUNX1*, high hyperdiploid
RGS12	4p16.3	Cell signaling	High hyperdiploid
BBC3	19q13.3–13.4	Apoptosis	*ETV6/RUNX1*
BOK	2q37.3	Apoptosis	*ETV6/RUNX1*, high hyperdiploid
CASP2	7q34-35	Apoptosis	High hyperdiploid
CASP9	1p36.1–36.3	Apoptosis	*ETV6/RUNX1*, high hyperdiploid
CD36	7q11.2	Apoptosis	High hyperdiploid
DBC1	9q32–33	Apoptosis	*ETV6/RUNX1*
FOXO3A	6q21	Apoptosis	High hyperdiploid
CD151	11p15.5	Cell adhesion	*ETV6/RUNX1*
CLDN1	3q28–29	Cell adhesion	High hyperdiploid
DMP1	4q21	Cell adhesion	High hyperdiploid
GNG4	1q42.3	Cell growth	*ETV6/RUNX1*
IGFBP6	12q13	Cell growth	High hyperdiploid
LTBP1	2p21–22	Cell growth	*ETV6/RUNX1*, high hyperdiploid
ARHGAP25	2p14	Signal transduction	High hyperdiploid
ARTN	1p32–33	Signal transduction	High hyperdiploid
CD2	1p13	Signal transduction	High hyperdiploid
CDK2AP1	12q24.31	Cell cycle	High hyperdiploid
PDPN	1p36.21	Cell cycle	*ETV6/RUNX1*, high hyperdiploid
BRCA1	17q21	DNA repair	High hyperdiploid
CDC14B	9q22.33	Protein dephosphorylation	*ETV6/RUNX1*
CES2	16q22.1	Metabolism	High hyperdiploid
CPSF3L	1p36.33	snRNA processing	*ETV6/RUNX1*, high hyperdiploid
FAT	4q35	Morphogenesis	*ETV6/RUNX1*
FGFBP1	4p15–16	Cell proliferation	High hyperdiploid
HIST1H1A	6p21.3	Nucleosome assembly	High hyperdiploid
PLGLB1	2p11.2	Unknown	High hyperdiploid
SMURF1	7q22.1	Cell differentiation	High hyperdiploid
TRIM10	6p21.3	Hematopoiesis	High hyperdiploid
UCHL1	4p14	Various	*ETV6/RUNX1*, high hyperdiploid

[a]Defined as a log2 difference of at least 0.4 compared with the mean expression in cases without a methylation peak.
[b]Function and location are based on data in NCBI Entrez Gene (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene). Only one major function per gene is given.
[c]As identified by the segmentation analysis.

level in cases with methylation peaks than in those without such peaks and 58 did not differ between methylated and non-methylated cases (Supplementary Material, Table S12). The remaining 36 (34%) genes showed a methylation peak associated with a significant decrease in expression (Table 3).

The genomic positions of the methylation peaks, the peak sizes, the probe densities and the positions of the peaks in relation to the transcription start sites of all genes listed in Tables 2 and 3 are given in Supplementary Material, Table S13.

DNA methylation hotspots throughout the genome

A total of 58 different regions/hotspots with a peak score density higher than four hypermethylated genes per 4 Mb was identified, with 30 of these being common to both *ETV6/RUNX1*-positive and high hyperdiploid ALLs. The *ETV6/RUNX1*-positive ALLs harbored eight unique methylation hotspots, whereas the high hyperdiploid ALLs had 20 unique methylation hotspots (Fig. 2, Supplementary Material, Fig. S3).

Clustering analysis of chromosome-wide methylomes and gene expression data

Unsupervised clustering and principal component analyses (PCA) of the bacterial artificial chromosome (BAC) array chromosome-wide methylome profiles could successfully subgroup the majority of *ETV6/RUNX1*-positive and high

Table 3. The 36 hypermethylated genes recurrently methylated and/or involved in B-cell development or mutated or rearranged in B-lineage ALL that displayed significantly[a] decreased expression in *ETV6/RUNX1*-positive and/or high hyperdiploid cases

Gene	Location[b]	Function[b]	Genetic subgroup	Inclusion criterion
BBC3	19q13.3–13.4	Apoptosis	*ETV6/RUNX1*, high hyperdiploid	Recurrently methylated
BCL10	1p22	Apoptosis	*ETV6/RUNX1*	Supplementary Material, Table S9
BOK	2q37.3	Apoptosis	*ETV6/RUNX1*, high hyperdiploid	Recurrently methylated
CASP9	1p36.1–36.3	Apoptosis	*ETV6/RUNX1*, high hyperdiploid	Recurrently methylated
DBC1	9q32–33	Apoptosis	*ETV6/RUNX1*, high hyperdiploid	Recurrently methylated+Supplementary Material, Table S2
FOXO3A	6q21	Apoptosis	*ETV6/RUNX1*, high hyperdiploid	Recurrently methylated
TP53I3	2p23.3	Apoptosis	*ETV6/RUNX1*, high hyperdiploid	Recurrently methylated
AR	Xq11.2–12	Cell signaling	high hyperdiploid	Supplementary Material, Table S1
CSF1R	5q33–35	Cell signaling	high hyperdiploid	Supplementary Material, Table S9
GALNAC4S-6ST	10q26	Cell signaling	*ETV6/RUNX1*, high hyperdiploid	Recurrently methylated
NOTCH1	9q34.3	Cell signaling	*ETV6/RUNX1*, high hyperdiploid	Supplementary Material, Table S9
PRLR	5p13–14	Cell signaling	*ETV6/RUNX1*, high hyperdiploid	Recurrently methylated+Supplementary Material, Table S1
SFRP1	8p11.1–12	Cell signaling	high hyperdiploid	Supplementary Material, Tables S1 and S2
TGFBR1	9q22	Cell signaling	high hyperdiploid	Recurrently methylated
CEBPA	19q13.1	Transcription regulation	high hyperdiploid	Supplementary Material, Table S10
HLF	17q22	Transcription regulation	*ETV6/RUNX1*, high hyperdiploid	Recurrently methylated+Supplementary Material, Table S10
MSC	8q21	Transcription regulation	*ETV6/RUNX1*, high hyperdiploid	Recurrently methylated
NFKB2	10q24	Transcription regulation	*ETV6/RUNX1*	Supplementary Material, Table S9
FAT	4q35	Morphogenesis	*ETV6/RUNX1*, high hyperdiploid	Recurrently methylated+Supplementary Material, Table S1
HCK	20q11–12	Morphogenesis	high hyperdiploid	Supplementary Material, Table S2
NEFL	8p21	Morphogenesis	*ETV6/RUNX1*, high hyperdiploid	Supplementary Material, Table S1
TPD52	8q21	Morphogenesis	*ETV6/RUNX1*, high hyperdiploid	Recurrently methylated
CCND2	12p13	Cell cycle	high hyperdiploid	Supplementary Material, Table S2
GAK	4p16	Cell cycle	*ETV6/RUNX1*, high hyperdiploid	Recurrently methylated
PRKCDBP	3p21.3	Cell cycle	*ETV6/RUNX1*, high hyperdiploid	Recurrently methylated+Supplementary Material, Table S1
IGF2	11p15.5	Cell proliferation	high hyperdiploid	Supplementary Material, Table S1
LDOC1	Xq27	Cell proliferation	*ETV6/RUNX1*, high hyperdiploid	Recurrently methylated
CD44	11p13	Cell adhesion	*ETV6/RUNX1*, high hyperdiploid	Supplementary Material, Tables S1 and S2
CDH13	16q24.2–24.3	Cell adhesion	high hyperdiploid	Supplementary Material, Tables S1 and S2
BRCA1	17q21	DNA repair	high hyperdiploid	Supplementary Material, Table S1
WRN	8p11.2–12	DNA repair	high hyperdiploid	Recurrently methylated+Supplementary Material, Table S1
CDC14B	9q22.3	Protein dephosphorylation	*ETV6/RUNX1*, high hyperdiploid	Recurrently methylated
RBP1	3q23	Transport	*ETV6/RUNX1*, high hyperdiploid	Supplementary Material, Table S1
SDHC	1q23.3	Electron transport	*ETV6/RUNX1*, high hyperdiploid	Recurrently methylated
TES	7q31.2	Ion binding	*ETV6/RUNX1*, high hyperdiploid	Recurrently methylated+Supplementary Material, Table S1
UCHL1	4p14	Various	*ETV6/RUNX1*, high hyperdiploid	Recurrently methylated

[a]Defined as a log2 difference of at least 0.4 compared with the mean expression in cases without a methylation peak.
[b]Function and location are based on data in NCBI Entrez Gene (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene). Only one major function per gene is given.

hyperdiploid cases (Fig. 3), indicating that they, to a large extent, are characterized by different chromosome-wide methylation patterns and expression profiles. In agreement with previous findings (13), gene rich BAC clones tended to be more methylated than those harboring no or only a few genes (Fig. 4A, Supplementary Material, Fig. S4A). However, it is important to stress that a high methylation level of a BAC clone not necessarily was associated with CpG hypermethylation of individual genes. In fact, a large proportion of the 71 different genes that had methylation peaks and that were underexpressed (Tables 2 and 3) did not map to hypermethylated BAC clones (data not shown).

Comparison of chromosome-wide methylomes in *ETV6/RUNX1*-positive and high hyperdiploid ALL

When comparing the mean methylation levels in the two ALL types, differences between the high hyperdiploid and the *ETV6/RUNX1*-positive ALL as regards chromosome-wide methylation patterns on chromosomes 6, 10, 14, 16, 17, 18, 19 and 21 are clearly seen, with the former cases generally displaying a lower methylation of these chromosomes (Fig. 4, Supplementary Material, Fig. S4). It is noteworthy that the majority of these chromosomes are the ones that are most commonly gained in ALL with high hyperdiploidy (1).

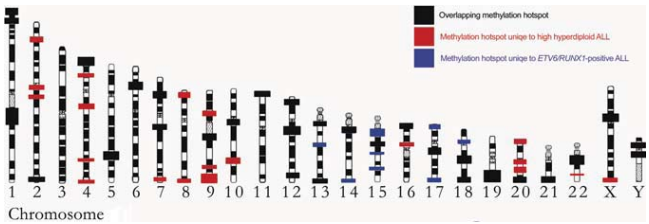


Figure 2. Fifty-eight different methylation hotspots, derived from the promoter CpG island array, were identified in high hyperdiploid and *ETV6/RUNX1*-positive ALLs. The hotspots common to both genetic subtypes are shown as black boxes ($n = 30$), the hotspots unique for the *ETV6/RUNX1*-positive cases are blue ($n = 8$), and the ones unique for high hyperdiploid ALLs are red ($n = 20$).

The only exceptions were chromosome 17, which did not display decreased methylation despite being trisomic in nine of the high hyperdiploid cases, and chromosomes 16 and 19, which were less methylated but only trisomic in one of the *ETV6/RUNX1*-positive cases. Furthermore, when comparing the mean methylation levels per chromosome within the high hyperdiploid group, they were lower in cases with trisomy/tetrasomy for chromosomes 4, 8, 10 and 21 than in cases with disomies of these chromosomes (Supplementary Material, Fig. S5). Thus, gains in high hyperdiploid cases were strongly associated with decreased methylation of the additional chromosomes. The same phenomenon was not seen in *ETV6/RUNX1*-positive ALL. Among the 10 cases analyzed, 16 had either trisomy 21 or gain of a large part of that chromosome (Table 1). However, there were no methylation differences between those with or without such gains (Supplementary Material, Fig. S5).

Gene expression profiling of disomic and trisomic/tetrasomic chromosomes

Lowess curves of mean-centered expression ratios of disomies, trisomies and tetrasomies for all chromosomes within the high hyperdiploid subgroup and between high hyperdiploid cases and all other pediatric B-lineage ALL analyzed in the study by Andersson *et al*. (19) were compared. We observed global overexpression of genes on the additional chromosomes in the high hyperdiploid cases, when compared with other ALL subtypes as well as when comparing disomic and trisomic/tetrasomic chromosomes within the high hyperdiploid group (Supplementary Material, Fig. S6).

Bisulfite sequencing

In the validation assay of the promoter CpG island methylation profiling, the array results corresponded to the bisulfite sequencing results of the seven selected genes (*BAD, BBC3, CAV1, CDK2AP1, NPM1, PRKCDBP* and *THEM4*) in all cases, i.e. low or moderate methylation was found when the investigated regions were situated outside the peak and hypermethylation was detected when the analyzed segments were situated within the peak region (Supplementary Material, Fig. S7).

DISCUSSION

During recent years, our knowledge about the global patterns of genomic alterations in childhood ALL has expanded dramatically (4,7,8). In contrast, no studies have focused on the genome-wide epigenetics of pediatric ALL; this is surprising considering that CpG hypermethylation of tumor suppressor genes has been demonstrated to be frequent in various neoplastic disorders (10). For this reason, we performed an in-depth analysis of global DNA methylation changes in the two most common genetic types of pediatric hematologic malignancies, including only primary patient material in order to avoid methylation artifacts arising from *in vitro* culturing.

It is noteworthy that none of the 30 genes displaying the highest methylation peak scores in the two genetic subgroups (Supplementary Material, Tables S5–S7) have previously been implicated as hypermethylated in neoplasia (Supplementary Material, Tables S1 and S2), a finding that emphasizes the value of performing genome-wide methylation analyses to identify novel hypermethylated candidate genes. As seen in Supplementary Material, Tables S5–S7, transcription regulation factors were particularly frequent among the genes with the highest methylation scores. Most likely, some of these genes, in particular those common to both subtypes (Supplementary Material, Table S6), may represent genes that are physiologically hypermethylated in early B-cell precursors and that hence may be associated with normal B-cell development. On the other hand, the genes that were differentially hypermethylated in the two different genetic subgroups (Supplementary Material, Table S7) are probably of leukemogenic importance.

Interestingly, among the genes with significant promoter methylation (Supplementary Material, Figs S1 and S2), only 30–60% were shown to be underexpressed (Tables 2 and 3). This clearly shows that hypermethylation per se does not necessarily lead to decreased gene expression. This has previously been demonstrated in high hyperdiploid ALL for the *FHIT* gene, which displayed no differential expression between methylation-positive and -negative cases (21). In fact, among the 16 hypermethylated genes identified in the study by Paulsson *et al*. (21), six were also found in the present study, namely *CD44, CDH13, FHIT, PAX6, RARB* and *WT1*, and expression data were available for five of these, showing downregulation of *CD44* and *CDH13*, no

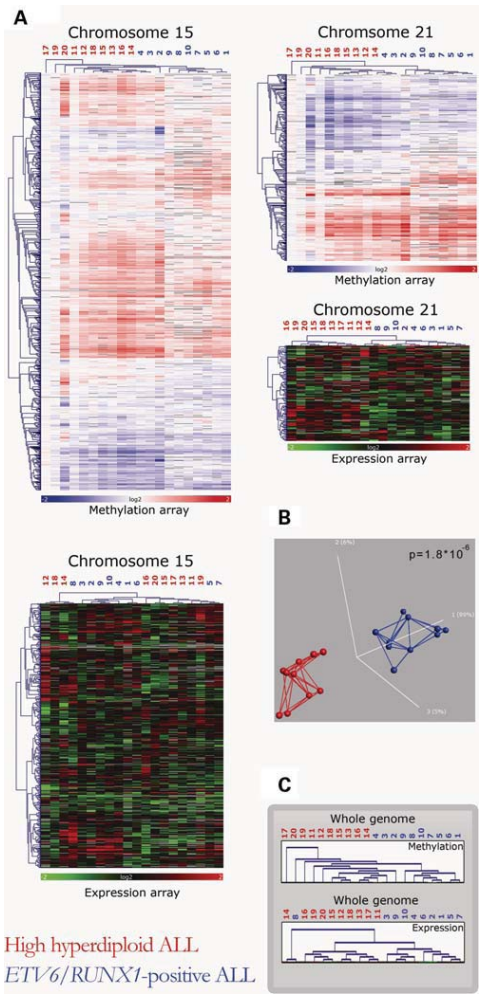


Figure 3. Analysis of the chromosome-wide BAC array methylome profiles could successfully subgroup the majority of *ETV6/RUNX1*-positive and high hyperdiploid ALLs, both on the whole genome level as well as on individual di- or trisomic chromosomes. (**A**) Unsupervised hierarchical clustering analysis of chromosomes 15 and 21. The former chromosome was disomic in all cases, whereas the latter was gained in all high hyperdiploid ALLs (red) and in the majority of the *ETV6/RUNX1*-positive cases (blue). (**B**) Principal component analysis of the chromosome wide methylation data. (**C**) Enhanced dendrograms of unsupervised clustering analysis of whole genome methylation and expression data.

differential expression for *FHIT* and *RARB*, and upregulation of *WT1*, again emphasizing that other factors besides CpG island methylation influence gene expression patterns. As regards the hypermethylated and underexpressed genes (Tables 2 and 3), the three most common GO groups were (i) transcription regulation, with *BRDT, HLF, SOX30* and *SSX4* being hypermethylated in the majority of cases, (ii) apoptosis, with *BBC3* and *FOXO3A* frequently being hypermethylated in both high hyperdiploid and *ETV6/RUNX1*-positive ALLs and (iii) cell signaling, with frequent hypermethylation of the *CSF2RA, GALNAC4S-6ST* and *PTPN3* genes (Supplementary Material, Figs S1 and S2). These genes could hence be potentially important targets and markers for future treatment trials with demethylating agents.

It has previously been demonstrated that simultaneous occurrence of hypermethylation and deletion causing biallelic gene inactivation is a rare event (22). Our findings agree well with this since none of the deletions identified by the array CGH analysis (Table 1) were associated with hypermethylated BAC clones or promoters in the retained chromosomal regions (Fig. 5 and data not shown). Cooperation between methylation events and copy number alterations is therefore unlikely to play a pathogenetically significant role in pediatric ALL.

The present genome-wide promoter methylation analysis identified a total of 58 different methylation hotspots, with eight and 20 being unique for the *ETV6/RUNX1*-positive and high hyperdiploid ALLs, respectively (Fig. 2, Supplementary Material, Fig. S3). One large methylation hotspot common to both subtypes was located on chromosome arm 19q (Fig. 2). Interestingly, this methylation hotspot has previously been reported in ALL and harbors several maternally imprinted genes (16). This could suggest that the other methylation hotspots identified herein also include imprinted genes. To test this possibility we used the comprehensive mapping information provided by Luedi *et al*. (23) on all known, or putatively, imprinted human genes. We then ascertained the number of imprinted genes per chromosome band and evaluated how often bands with a specific number of imprinted genes harbored methylation hotspots. As seen in Figure 6, there is a high correlation between the presence of imprinted genes and methylation hotspots, with the frequencies of methylation hotspots increasing with the number of imprinted genes. This strongly suggests that regions rich in imprinted genes are associated with methylation hotspots in *ETV6/RUNX1*-positive and high hyperdiploid ALLs. One may hypothesize that imprinted genes are more susceptible to *de novo* methylation of the active allele and hence biallelic silencing.

The detection of subtype-specific methylation hotspots (Fig. 2) indicates that *ETV6/RUNX1*-positive and high hyperdiploid ALLs are characterized by different global methylation patterns. Further support for this was derived from the clustering analyses which successfully could subgroup the *ETV6/RUNX1*-positive and high hyperdiploid cases (Fig. 3). It has previously been reported that the profiles of CpG island hypermethylation of tumor suppressor genes differ between different tumor entities (10) and that acute myeloid leukemia and ALL can be separated using a 57 gene/249 CpG methylation array (17). The present study is the first to use global methylation patterns to distinguish different genetic subtypes of ALL. Thus, genome-wide methylation analyses can be used to subdivide morphologically identical, but genetically and clinically distinct, hematologic malignancies, a finding that may have ramification also for other tumor types.

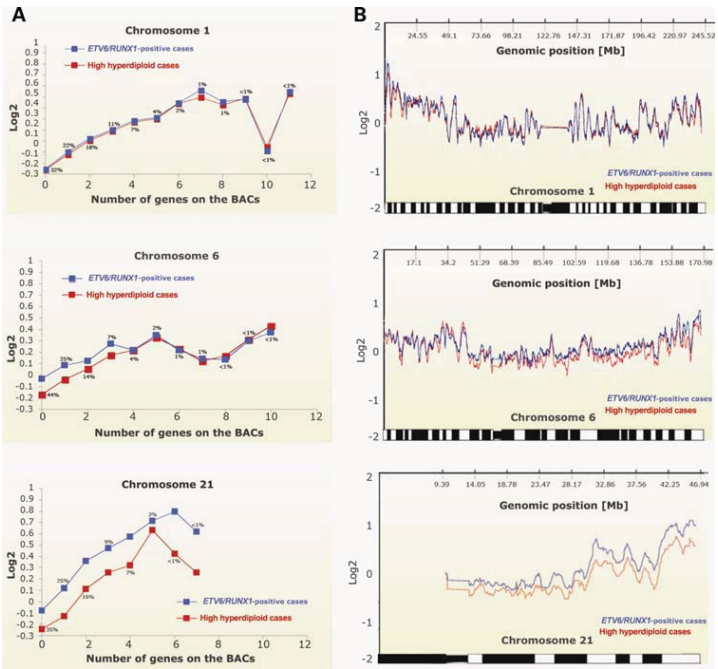


Figure 4. (**A**) Mean log2 ratios of the high hyperdiploid (red) and *ETV6/RUNX1*-positive (blue) ALLs plotted against the number of genes on the BACs for chromosomes 1, 6 and 21. Chromosome 1 was disomic in all cases, chromosome 6 was trisomic in all but one of the high hyperdiploid cases and disomic in all *ETV6/RUNX1*-positive ALLs and chromosome 21 was gained in the vast majority of all cases. The percentages that the BAC clones with different number of genes comprise of the whole chromosome are given above each BAC group. As seen, the high hyperdiploid cases display a lower methylation of the frequently gained chromosomes 6 and 21 when compared with the *ETV6/RUNX1*-positive cases. (**B**) Mean log2 ratios of the high hyperdiploid (red) and *ETV6/RUNX1*-positive (blue) ALLs plotted against genomic positions on chromosomes 1, 6 and 21. Again, the same pattern, i.e. lower methylation of gained chromosomes, is seen for both chromosomes 6 and 21.

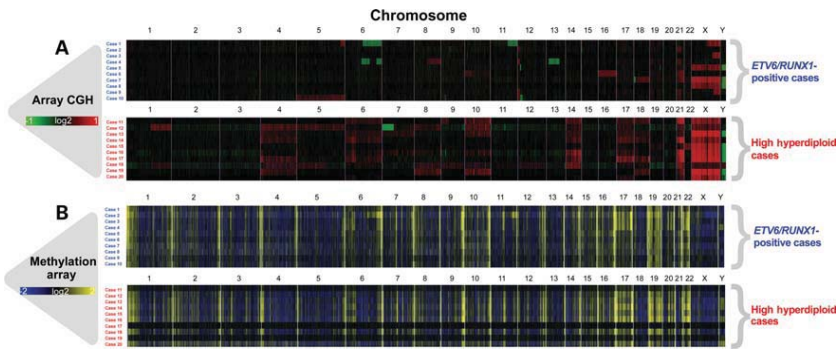


Figure 5. Genome and DNA methylome profiles of *ETV6/RUNX1*-positive and high hyperdiploid ALLs. (**A**) Heat maps displaying genomic imbalances detected by array CGH. The gained chromosomes in the high hyperdiploid cases are clearly seen (indicated in red). The imbalance patterns of the *ETV6/RUNX1*-positive cases are quite different, with more losses (indicated in green), for example deletions of 6q. (**B**) Heat maps displaying chromosome-wide methylation profiles. As seen, none of the genomic imbalances identified by array CGH were associated with hypermethylated BAC clones in the corresponding regions.

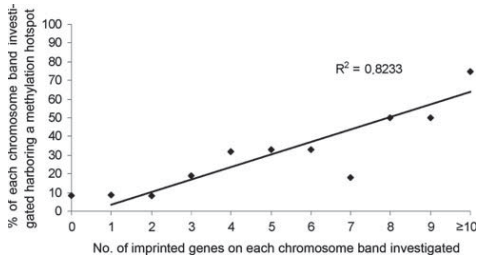


Figure 6. Mapping information on all known, or putatively, imprinted human genes was retrieved from Luedi *et al*. (23) and compared with the location of methylation hotspots. As seen, there is a high correlation between the presence of imprinted genes and hotspots, with the frequencies of hotspots increasing significantly with the number of imprinted genes.

An intriguing and unexpected finding was the clear-cut differences between the high hyperdiploid and the *ETV6/RUNX1*-positive ALLs as regards the chromosome-wide methylation patterns, with the majority of the commonly gained chromosomes in the former subtype being less methylated than their disomic counterparts (Fig. 4, Supplementary Material, Fig. S4). Since we did not observe the same phenomenon in the *ETV6/RUNX1*-positive ALL cases with or without trisomy 21 (Supplementary Material, Fig. S5) it is apparent that gains as such do not correlate with chromosome-wide differential methylation and that the methylation changes hence depend on other factors, for example chromosome type, number of gains and cell of origin. We know of only one previous study reporting similar findings, namely the one by Weber *et al.* (13) who identified an overall lowered methylation in gene-poor regions of trisomic chromosomes in a colon cancer cell line. However, we observed decreased methylation in primarily gene-poor regions only on chromosomes 6 and 8 in the high hyperdiploid ALLs. In contrast, this was not the case for chromosomes 4, 10, 14, 16, 18, 19 and 21; on these chromosomes the lowered methylation was equally distributed irrespective of the number of genes on the BAC clones (Supplementary Material, Figs S4 and S5). The latter finding agrees well with previous data showing no clear bias for decreased methylation only of gene-poor regions in the inactive and hypomethylated X chromosome in females (24).

It has previously been demonstrated that aneuploidy in yeast is associated with expression of the great majority of the genes on the extra chromosomes. However, even if the amount of transcript is increased, the protein levels are often not increased because of either non-translation or degradation (25). Thus, the overexpression of genes on additional chromosomes in high hyperdiploid cases, when compared with other ALL subtypes (19) or when comparing disomic and trisomic chromosomes within the high hyperdiploid group (Supplementary Material, Fig. S6), does not by necessity translate into protein abundance. Hence, the role of the decreased methylation of the gained chromosomes could be unrelated to gene expression. Instead, the differential methylation may be necessary for proper compartmentalization of the tri-/tetrasomic chromosomes in interphase, akin to Barr body formation of the inactive X chromosome, which, surprisingly, has been shown to be globally hypomethylated relative to its active counterpart (13).

Recently, it was shown that aneuploidy generally inhibits tumorigenesis in murine experimental systems, being associated with decreased proliferation; however, the opposite effect—oncogenesis—may occur with gains of certain specific chromosomes (26,27). An abnormal chromosome number as such is hence unlikely to be sufficient to drive oncogenesis—other genetic or epigenetic changes are needed in addition. The lowered methylation identified herein may be one such event, perhaps abrogating the reduced cellular fitness associated with aneuploidy, and may have ramifications not only for the pathogenesis of high hyperdiploid ALL but also for other disorders with numerical chromosome anomalies.

MATERIALS AND METHODS

Patients and samples

Samples from 20 children with B-cell precursor ALL, comprising 10 with high hyperdiploidy and 10 with the *ETV6/RUNX1* fusion, were included in the study (Table 1). All patients had been diagnosed and treated according to the NOPHO-ALL 1992 or 2000 protocols at the Departments of Pediatric Oncology and Hematology, Lund University Hospital, Lund, and Linköping University Hospital, Linköping, Sweden. The study was reviewed and approved by the regional research ethics boards at Lund and Linköping Universities, and informed consent was obtained from the patients' parents or guardians in accordance with the Declaration of Helsinki. DNA was extracted, at the time of diagnosis, from bone marrow samples in cases 1–19 and from peripheral blood in case 20. All the samples were cytogenetically characterized by conventional chromosome banding analysis. The presence of the cytogenetically cryptic *ETV6/RUNX1* fusion was ascertained by RT–PCR and verified by FISH using locus-specific probes for the *ETV6* and *RUNX1* genes. The additional chromosomes identified by G-banding in the high hyperdiploid cases were, in the present study, all confirmed by array CGH (see in what follows). DNA and RNA were extracted using standard methods, and the DNA was purified further with the DNeasy Blood & Tissue Kit (Qiagen, Solna, Sweden) and quantified using a Nanodrop Spectrophotometer (Nanodrop Technologies, Wilmington, DE).

Array comparative genomic hybridization

Array CGH slides, containing 32 433 tiling BAC clones covering at least 98% of the human genome, were produced at the SCIBLU DNA microarray resource center at Lund University, Sweden. Labeling of DNA, slide preparation and hybridization were performed as described previously (20). Male genomic DNA (Promega, Madison, WI) was used as reference in all hybridizations.

Analyses of the microarray images were performed as described previously (8). To identify imbalances the BASE (28), the CGH Explorer software (29) and the TM4 microarray software suite (30) were applied. Classification as gain or loss was based on the log2 ratios, where $> \pm 0.5$ in five adjacent clones were classified as aberrant, with ± 0.5–1.0

interpreted as duplications/hemizygous deletions and $> \pm 1.0$ as amplifications/homozygous deletions.

Methylated DNA immunoprecipitation

Six micrograms of DNA from each sample was diluted in deionized water, yielding a concentration of 75 ng/μl. The DNA was sonicated with a Sonifier set to 20% amplitude for 0.5 s in 60 cycles to generate fragments ranging from 200 to 1000 bp (verified by gel electrophoresis). The fragmented DNA was diluted to a concentration of 13.4 ng/μl and denatured for 10 min at 95°C, after which an input fraction (~250 ng), serving as a reference in subsequent microarray hybridizations, was collected and stored at −80°C. The remaining DNA was immunoprecipitated overnight with a monoclonal antibody against 5-methylcytidine (Diagenode, Liège, Belgium) in a total of 500 μl 1× IP buffer at 4°C, as previously described (13). The mixture was then incubated with 80 μl of Protein A Agarose beads (Invitrogen, Carlsbad, CA) for 2 h at 4°C and subsequently washed three times with 1× IP buffer to remove DNA with no or low level methylation. The beads were then treated overnight with proteinase K at 55°C. After centrifugation, methylated DNA in the supernatant was recovered using standard phenol–chloroform extraction and ethanol precipitation. Both the input fraction and the methylated DNA were then amplified using the WGA2 kit (Sigma-Aldrich, St. Louis, MO), generating sufficient amounts (~6 μg) for the various microarray hybridizations, as detailed in what follows.

CpG island and promoter methylation profiling

Four micrograms of methylated and input fraction DNA were analyzed using the NimbleGen CpG Island-Plus-Promoter Array (Roche NimbleGen, Madison, WI) by the NimbleGen Service Group (Reykjavik, Iceland) who performed the quality control, labeling, hybridization and scanning. This array platform covers all 28 226 UCSC-annotated CpG islands and promoter regions for all RefSeq genes, including 385 000 probes of 50–75 mer length per array, with an upstream and downstream promoter tiling of 800 bp and 200 bp, respectively.

Signal intensity data were extracted from the scanned images. Each feature on the array has a corresponding log2 ratio, which is the ratio of the input signals for the methylated DNA and the input fraction DNA. The log2 ratio was scaled in order to center the ratio data around zero by subtracting the bi-weight mean for the log2 ratio values for all features on the array from each log2 ratio value. A modified ACME algorithm (31), where a fixed-length window (750 bp) is placed around each consecutive probe, and the one-sided Kolmogorov–Smirnov test were then applied on the scaled log2 ratio data to determine whether the probes were drawn from a significantly more positive distribution (peak) of intensity log2 ratios than the other probes in the array. The resulting score for each probe is the $-\log10$ *P*-value from the windowed Kolmogorov–Smirnov test around that probe.

Delineation of hypermethylated genes

Peak data were generated by searching for probes above a *P*-value minimum cut-off ($-\log10$) of 2. Peaks within 500 bp of each other were merged. Each annotated gene was subsequently searched for peaks appearing in a specified promoter region around the transcription start site. The region searched was design-specific, spanning from 5 kb upstream to 1 kb downstream of the transcription start site. Genes that had at least two probes with peaks above the *P*-value minimum cut-off were considered to have significant CpG island methylation scores for hypermethylation. In this fashion only 5′ promoter CpG islands were investigated. The detected hypermethylated genes were analyzed in three different ways.

First, all genes with significant CpG island methylation scores were grouped according to ALL subtype in order to identify those that were uniquely hypermethylated in either *ETV6/RUNX1*-positive or high hyperdiploid cases. Furthermore, the chromosomal distribution of all genes was ascertained.

Second, an edge preserving smoother analysis (29), determining a sequence with as few jumps as possible, was performed. We used the Potts filter with the penalty parameter set to 2 and limited the least allowed aberration size to two clones. This segmentation analysis of the peak score data was performed in order to identify peaked genes throughout the genome that were highly enriched for methylation.

Third, we identified all genes that displayed methylation peaks in ≥30% of cases in either genetic ALL subgroup or that were present in Supplementary Material, Tables S1, S2 or S9–S11, which comprise genes known to be hypermethylated in malignant disorders, involved in B-cell development or mutated or rearranged in B-lineage ALL. Each gene was then functionally annotated using the EASE software (32) and correlated with gene expression data.

Identification of DNA methylation hotspots

Identification of regions with a dense occurrence of hypermethylated genes, i.e. genomic sites harboring a significantly higher number of peak-scored probes compared with the rest of the genome, was performed by plotting the spatial and number distribution of peak scored promoters using an Epanechenikov kernel density estimator in CGH Explorer (29). The size determinant was set to 4 Mb. In order to designate a chromosomal segment as a methylation hotspot at least four peak-scored probes had to be present within the 4 Mb segment analyzed.

Chromosome-wide methylation profiling

Of the methylated and input fraction DNA, 1.5 μg was differentially labeled with Cy3 and Cy5, respectively, and hybridized to 32K BAC array slides. Labeling of DNA, slide preparation, hybridization and analysis were performed as described previously (8,20). Classification as relative enrichment for methylation was based on the log2 ratio for each of the 32 433 BAC clones smoothed with the factor of 0.33, with a positive ratio indicating enrichment of methylation.

Unsupervised clustering analysis, after using a variance filter with a cut-off value of 0.25 standard deviation to limit the number of probes and a Pearson uncentered test (30) for average linkage clustering, was used to investigate methylation differences/similarities between the *ETV6/RUNX1*-positive and high hyperdiploid cases. In addition, a two

group comparison was performed using PCA in the Qlucore Software (Qlucore AB, Lund, Sweden).

Average log2 ratios for each BAC clone of the ten high hyperdiploid and the ten *ETV6/RUNX1*-positive ALLs, respectively, were plotted towards the genomic position on each chromosome. The two groups were also analyzed in relation to the gene density on each clone by ascertaining, using a custom-made script, the number of annotated RefSeq genes on each consecutive clone. All clones on each chromosome were grouped in relation to the number of genes they harbored, generally ranging from 0 to 12 genes. The average log2 ratio for each such group was then plotted against their gene content.

Global gene expression analysis

RNA extraction, amplification, labeling, hybridization to 27K cDNA microarray slides, scanning, post-hybridization washing and feature analysis have previously been described (19,33). In short, the samples were hybridized to microarray slides containing 25 648 cDNA clones (SCIBLU DNA microarray resource center) and amplified RNA from the Universal Human Reference (Agilent Technologies, Palo Alto, CA) was used as reference. The data analyses were all performed using the BASE (28), TM4 (30) and Statistica (StatSoft Scandinavia AB, Uppsala, Sweden) software. The same type of clustering analysis that was performed on the chromosome-wide methylation profiles (discussed earlier) was used on the expression data.

Bisulfite sequencing

To validate the reliability of the promoter array, eight genes (*BAD, BBC3, CAV1, CDK2AP1, NPM1, PRKCDBP, TES* and *THEM4*) were selected from the identified 167 recurring target genes (Supplementary Material, Fig. S2; see Results). DNA (2 μg) from eight leukemic samples, comprising four ALL with high hyperdiploidy and four with *ETV6/RUNX1*, was treated with sodium bisulfite using the EpiTect kit (Qiagen, Valencia, CA). The bisulfite-treated DNA was used as a template in standard PCR amplification using primers directed towards the promoter regions harboring the hypermethylated CpG islands, as indicated by peaks in the array (primer sequences available upon request). The primers were designed using the MethPrimer software (34). The PCR products were sub-cloned using the TOPO-TA system (Invitrogen, Carlsbad, CA) and sequenced using standard methods (MCLab, San Fransisco, CA). After sequencing 8–12 clones per gene and case, one gene (*TES*) was excluded because it harbored too few informative CpG sites. The sequences were analyzed using the BiQ Analyzer software (35). For a gene to be considered hypermethylated, the following criterion was used: at least 80% of the investigated clones should be methylated in at least 80% of the sequenced CpG sites.

Array data

The microarray data are deposited in the Array Express database (http://www.ebi.ac.uk/microarray.as/ae/).

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

Conflict of Interest statement: None declared.

FUNDING

This work was supported by grants from the Swedish Childhood Cancer Foundation; the Swedish Cancer Society; the Swedish Research Council and the IngaBritt and Arne Lundberg Foundation.

REFERENCES

1. Johansson, B., Mertens, F. and Mitelman, F. (2004) Clinical and biological importance of cytogenetic abnormalities in childhood and adult acute lymphoblastic leukemia. *Ann. Med.*, **36**, 492–503.
2. Greaves, M. (2005) In utero origins of childhood leukaemia. *Early Hum. Dev.*, **81**, 123–129.
3. Raynaud, S., Cavé, H., Baens, M., Bastard, C., Cacheux, V., Grosgeorge, J., Guidal-Giroux, C., Guo, C., Vilmer, E., Marynen, P. *et al.* (1996) The 12;21 translocation involving *TEL* and deletion of the other *TEL* allele: Two frequently associated alterations found in childhood acute lymphoblastic leukemia. *Blood*, **87**, 2891–2899.
4. Mullighan, C.G., Goorha, S., Radtke, I., Miller, C.B., Coustan-Smith, E., Dalton, J.D., Girtman, K., Mathew, S., Ma, J., Pounds, S.B. *et al.* (2007) Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature*, **446**, 758–764.
5. Paulsson, K., Horvat, A., Strömbeck, B., Nilsson, F., Heldrup, J., Behrendtz, M., Forestier, E., Andersson, A., Fioretos, T. and Johansson, B. (2008) Mutations of *FLT3*, *NRAS*, *KRAS*, and *PTPN11* are frequent and possibly mutually exclusive in high hyperdiploid childhood acute lymphoblastic leukemia. *Genes Chromosomes Cancer*, **47**, 26–33.
6. Forestier, E., Andersen, M.K., Autio, K., Blennow, E., Borgström, G., Golovleva, I., Heim, S., Heinonen, K., Hovland, R., Johannsson, J.H. *et al.* (2007) Cytogenetic patterns in *ETV6/RUNX1*-positive pediatric B-cell precursor acute lymphoblastic leukemia: A Nordic series of 245 cases and review of the literature. *Genes Chromosomes Cancer*, **46**, 440–450.
7. Lilljebjörn, H., Heidenblad, M., Nilsson, B., Lassen, C., Horvat, A., Heldrup, J., Behrendtz, M., Johansson, B., Andersson, A. and Fioretos, T. (2007) Combined high-resolution array-based comparative genomic hybridization and expression profiling of *ETV6/RUNX1*-positive acute lymphoblastic leukemias reveal a high incidence of cryptic Xq duplications and identify several putative target genes within the commonly gained region. *Leukemia*, **21**, 2137–2144.
8. Davidsson, J., Andersson, A., Paulsson, K., Heidenblad, M., Isaksson, M., Borg, Å., Heldrup, J., Behrendtz, M., Panagopoulos, I., Fioretos, T. *et al.* (2007) Tiling resolution array comparative genomic hybridization, expression and methylation analyses of dup(1q) in Burkitt lymphomas and pediatric high hyperdiploid acute lymphoblastic leukemias reveal clustered near-centromeric breakpoints and overexpression of genes in 1q22-32.3. *Hum. Mol. Genet.*, **16**, 2215–2225.
9. Conrad, B. and Antonarakis, S.E. (2007) Gene duplication: A drive for phenotypic diversity and cause of human disease. *Annu. Rev. Genomics Hum. Genet.*, **8**, 17–35.
10. Esteller, M. (2007) Epigenetic gene silencing in cancer: The DNA hypermethylome. *Hum. Mol. Genet.*, **16**, R50–R59.
11. Esteller, M. (2007) Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat. Rev. Genet.*, **8**, 286–298.
12. Fraga, M.F. and Esteller, M. (2002) DNA methylation: A profile of methods and applications. *Biotechniques*, **33**, 632–649.
13. Weber, M., Davies, J.J., Wittig, D., Oakeley, E.J., Haase, M., Lam, W.L. and Schübeler, D. (2005) Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat. Genet.*, **37**, 853–862.
14. Costello, J.F., Frühwald, M.C., Smiraglia, D.J., Rush, L.J., Robertson, G.P., Gao, X., Wright, F.A., Feramisco, J.D., Peltomäki, P., Lang, J.C.

et al. (2000) Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. *Nat. Genet.*, **24**, 132–138.
15. Keshet, I., Schlesinger, Y., Farkash, S., Rand, E., Hecht, M., Segal, E., Pikarski, E., Young, R.A., Niveleau, A., Cedar, H. *et al.* (2006) Evidence for an instructive mechanism of *de novo* methylation in cancer cells. *Nat. Genet.*, **38**, 149–153.
16. Taylor, K.H., Pena-Hernandez, K.E., Davis, J.W., Arthur, G.L., Duff, D.J., Shi, H., Rahmatpanah, F.B., Sjahputera, O. and Caldwell, C.W. (2007) Large-scale CpG methylation analysis identifies novel candidate genes and reveals methylation hotspots in acute lymphoblastic leukemia. *Cancer Res.*, **67**, 2617–2625.
17. Scholz, C., Nimmrich, I., Burger, M., Becker, E., Dörken, B., Ludwig, W-D. and Maier, S. (2005) Distinction of acute lymphoblastic leukemia from acute myeloid leukemia through microarray-based DNA methylation analysis. *Ann. Hematol.*, **84**, 236–244.
18. Suzuki, H., Gabrielson, E., Chen, W., Anbazhagan, R., van Engeland, M., Weijenberg, M.P., Herman, J.G. and Baylin, S.B. (2002) A genomic screen for genes upregulated by demethylation and histone deacetylase inhibition in human colorectal cancer. *Nat. Genet.*, **31**, 141–149.
19. Andersson, A., Olofsson, T., Lindgren, D., Nilsson, B., Ritz, C., Edén, P., Lassen, C., Råde, J., Fontes, M., Mörse, H. *et al.* (2005) Molecular signatures in childhood acute leukemia and their correlations to expression patterns in normal hematopoietic subpopulations. *Proc. Natl Acad. Sci. USA*, **102**, 19069–19074.
20. Jönsson, G., Staaf, J., Olsson, E., Heidenblad, M., Vallon-Christersson, J., Osoegawa, K., de Jong, P., Oredsson, S., Ringnér, M., Höglund, M. *et al.* (2007) High-resolution genomic profiles of breast cancer cell lines assessed by tiling BAC array comparative genomic hybridization. *Genes Chromosomes Cancer*, **46**, 543–558.
21. Paulsson, K., An, Q., Moorman, A.V., Parker, H., Molloy, G., Davies, T., Griffiths, M., Ross, F.M., Irving, J., Harrison, C.J. *et al.* (2009) Methylation of tumour suppressor gene promoters in the presence and absence of transcriptional silencing in high hyperdiploid acute lymphoblastic leukaemia. *Br. J. Haematol.*, **144**, 838–847.
22. Zardo, G., Tiirikainen, M.I., Hong, C., Misra, A., Feuerstein, B.G., Volik, S., Collins, C.C., Lamborn, K.R., Bollen, A., Pinkel, D. *et al.* (2002) Integrated genomic and epigenomic analyses pinpoint biallelic gene inactivation in tumors. *Nat. Genet.*, **32**, 453–458.
23. Luedi, P.P., Dietrich, F.S., Weidman, J.R., Bosko, J.M., Jirtle, R.L. and Hartemink, A.J. (2007) Computational and experimental identification of novel human imprinted genes. *Genome Res.*, **17**, 1723–1730.
24. Hellman, A. and Chess, A. (2007) Gene body-specific methylation on the active X chromosome. *Science*, **315**, 1141–1143.
25. Torres, E.M., Sokolsky, T., Tucker, C.M., Chan, L.Y., Boselli, M., Dunham, M.J. and Amon, A. (2007) Effects of aneuploidy on cellular physiology and cell division in haploid yeast. *Science*, **317**, 916–924.
26. Weaver, B.A.A., Silk, A.D., Montagna, C., Verdier-Pinard, P. and Cleveland, D.W. (2007) Aneuploidy acts both oncogenically and as a tumor suppressor. *Cancer Cell*, **11**, 25–36.
27. Williams, B.R., Prabhu, V.R., Hunter, K.E., Glazier, C.M., Whittaker, C.A., Housman, D.E. and Amon, A. (2008) Aneuploidy affects proliferation and spontaneous immortalization in mammalian cells. *Science*, **522**, 703–709.
28. Saal, L.H., Troein, C., Vallon-Christersson, J., Gruvberger, S., Borg, Å. and Peterson, C. (2002) BioArray Software Environment (BASE): a platform for comprehensive management and analysis of microarray data. *Genome Biol.*, **3**, Software/0003.1-0003.6.
29. Lingjærde, O.C., Baumbusch, L.O., Liestøl, K., Glad, I.K. and Børresen-Dale, L. (2005) CGH-Explorer: A program for analysis of array-CGH data. *Bioinformatics*, **21**, 821–822.
30. Saeed, A.I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., Braisted, J., Klapa, M., Currier, T., Thiagarajan, M. *et al.* (2003) TM4: A free, open-source system for microarray data management and analysis. *Biotechniques*, **34**, 374–378.
31. Scacheri, P.C., Crawford, G.E. and Davis, S. (2006) Statistics for ChIP-chip and DNase hypersensitivity experiments on NimbleGen arrays. *Methods Enzymol.*, **411**, 270–282.
32. Hosack, D.A., Dennis, G. Jr, Sherman, B.T., Lane, H.C. and Lempicki, R.A. (2003) Identifying biological themes within lists of genes with EASE. *Genome Biol.*, **4**, R70.
33. Andersson, A., Edén, P., Lindgren, D., Nilsson, J., Lassen, C., Heldrup, J., Fontes, M., Borg, Å., Mitelman, F., Johansson, B. *et al.* (2005) Gene expression profiling of leukemic cell lines reveals conserved molecular signatures among subtypes with specific genetic aberrations. *Leukemia*, **19**, 1042–1050.
34. Li, L.C. and Dahiya, R. (2002) MethPrimer: Designing primers for methylation PCRs. *Bioinformatics*, **18**, 1427–1431.
35. Bock, C., Reither, S., Mikeska, T., Paulsen, M., Walter, J. and Lengauer, T. (2005) BiQ Analyzer: Visualization and quality control for DNA methylation data from bisulfite sequencing. *Bioinformatics*, **21**, 4067–4068.

Supplementary material II

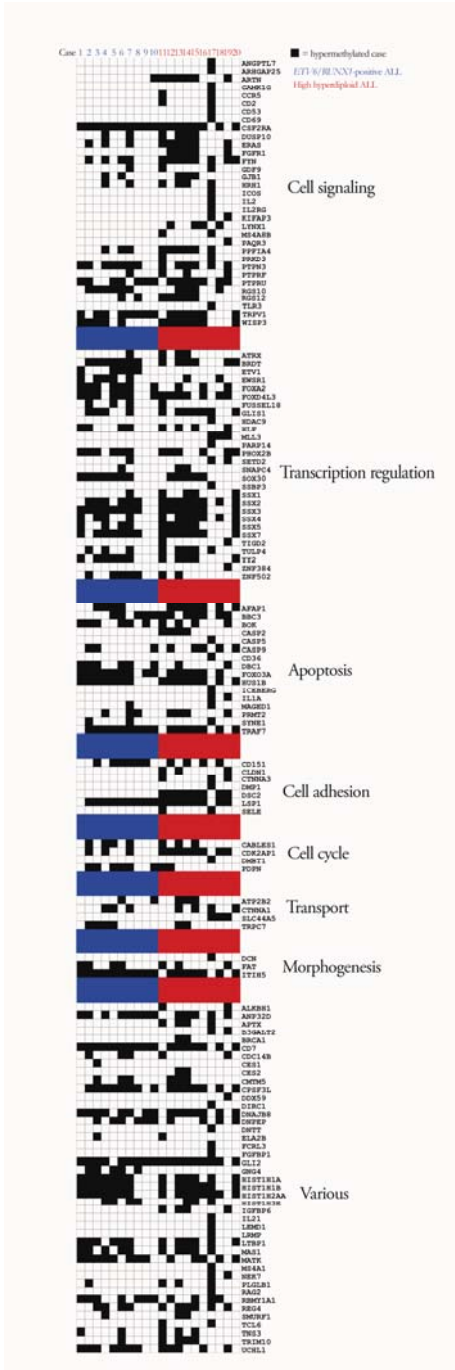


Figure S1. A total of 138 genes were identified by the segmentation analysis to be highly enriched for methylation among the *ETV6/RUNX1*-positive and high hyperdiploid ALLs.

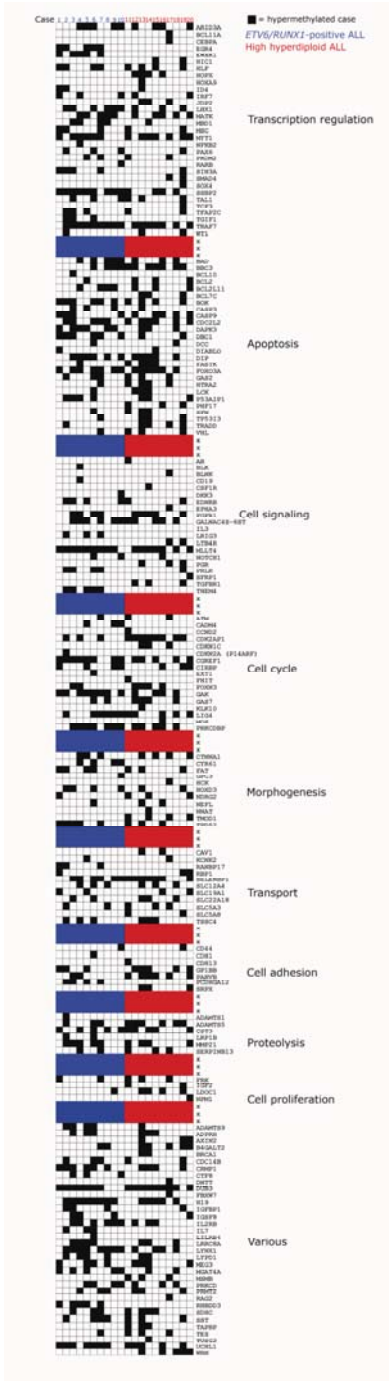
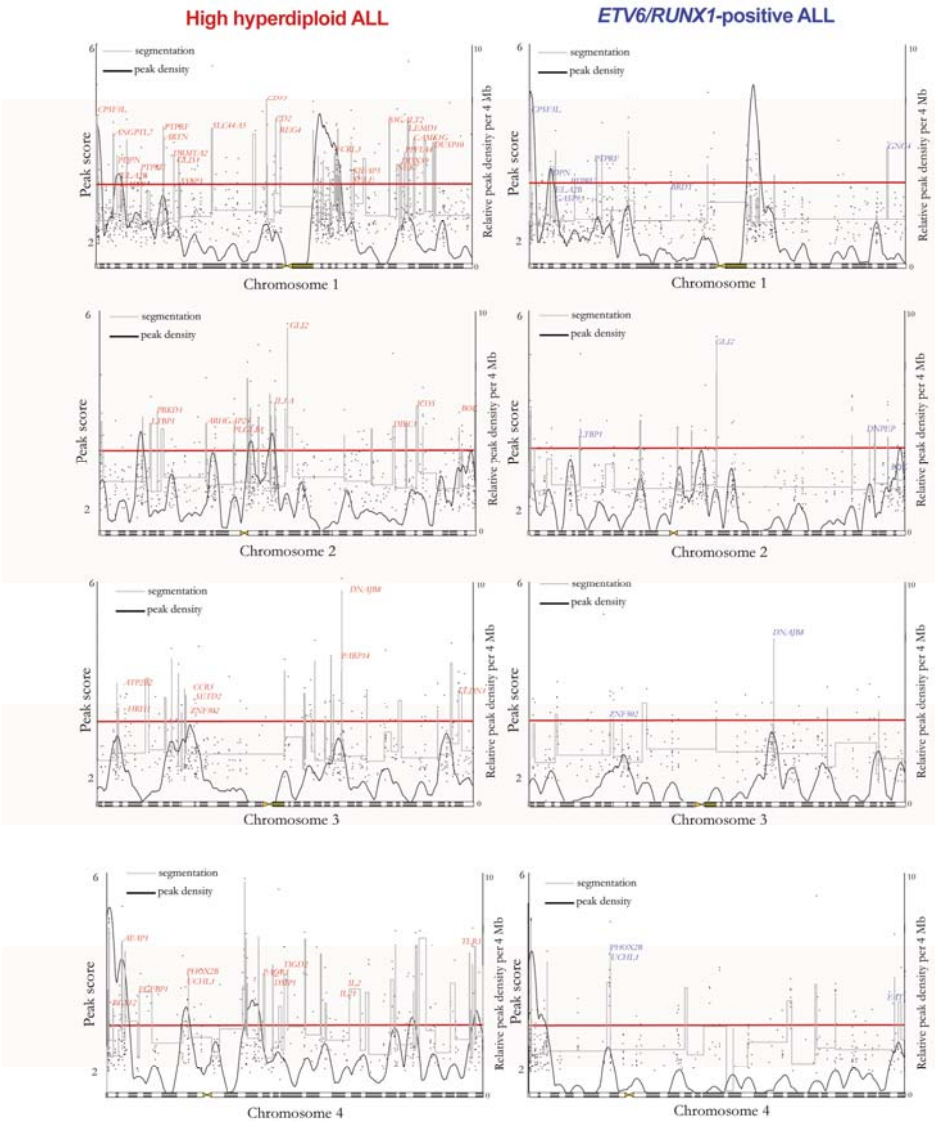
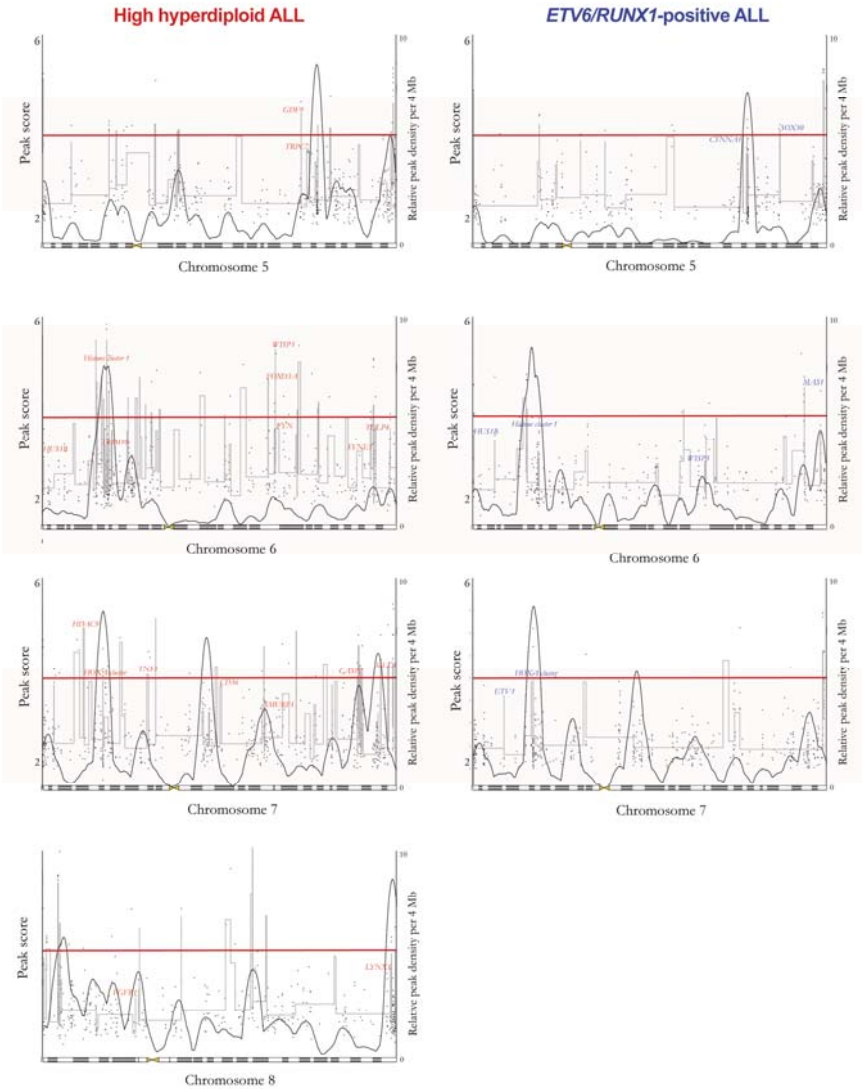
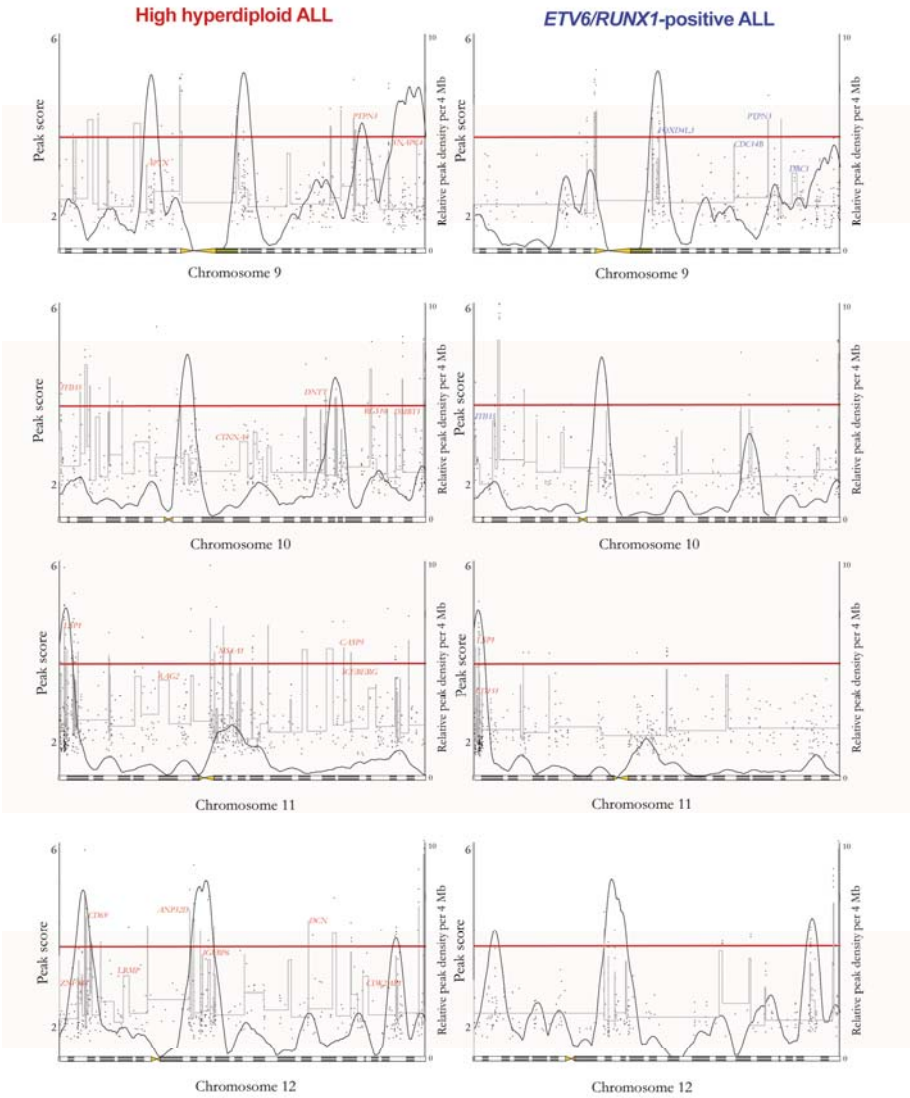
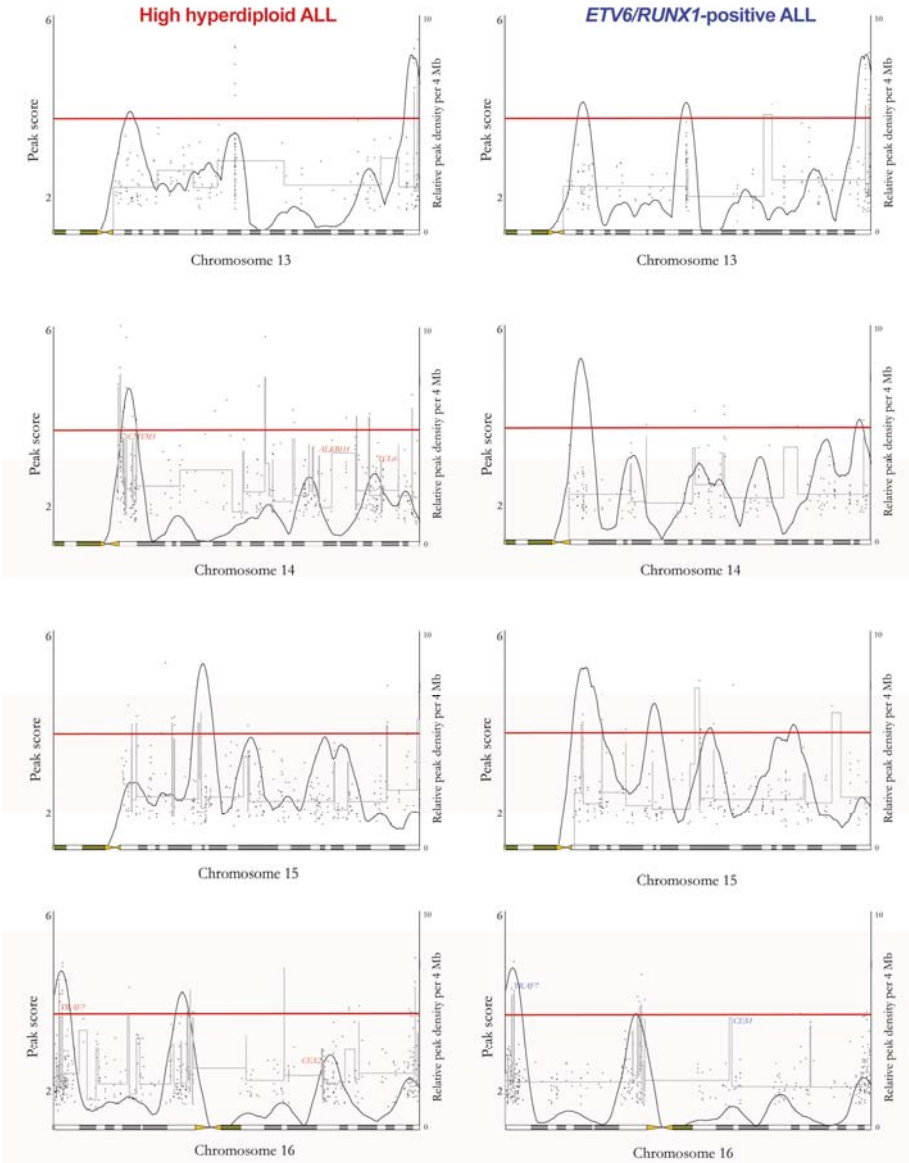


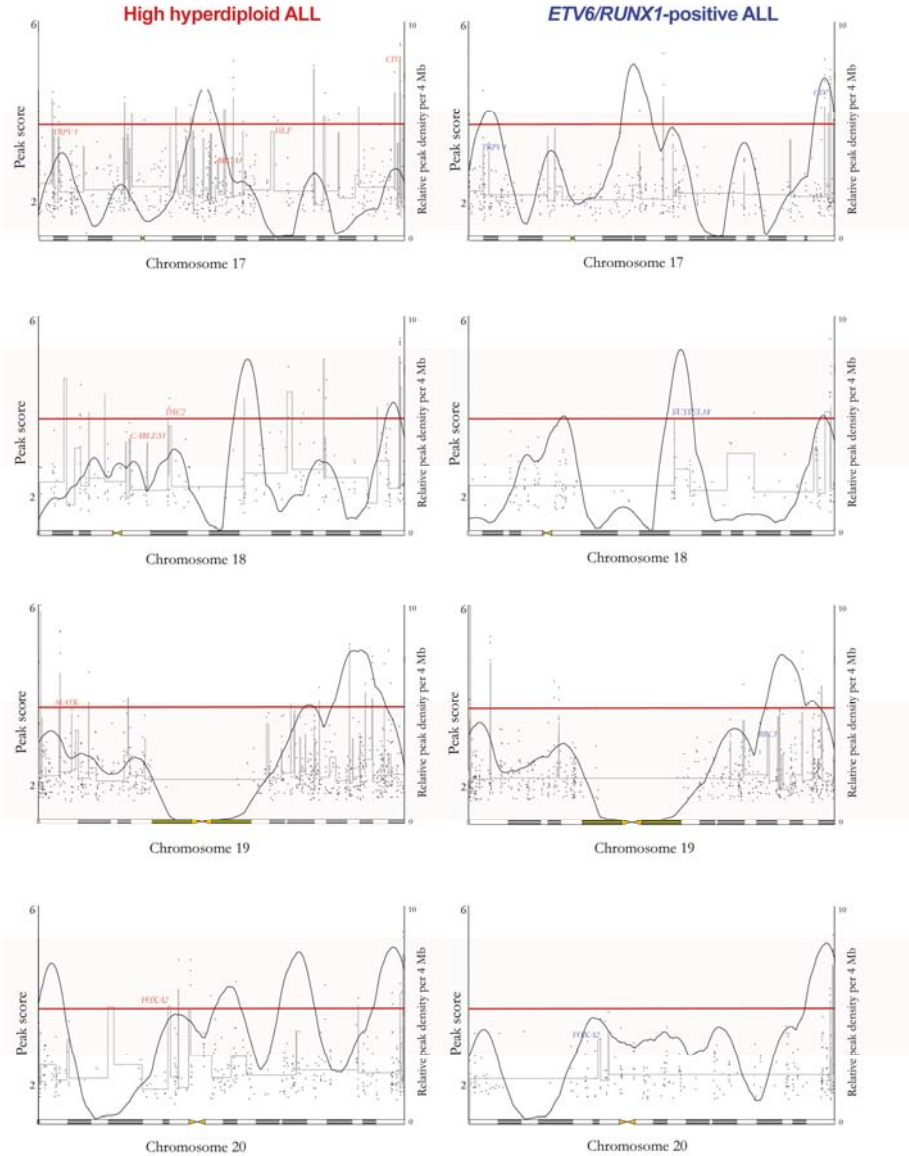
Figure S2. A total of 167 genes displaying methylation peaks in at least 30% of the cases in each genetic subgroup or being present in Tables S1, S2 and S9-S11 were identified.


High hyperdiploid ALL
ETV6/RUNX1-positive ALL
segmentation
peak density
Peak score
Relative peak density per 4 Mb
Chromosome 1
Chromosome 2
Chromosome 3
Chromosome 4


High hyperdiploid ALL
ETV6/RUNX1-positive ALL
Peak score
Relative peak density per 4 Mb
Chromosome 5
Chromosome 5
Chromosome 6
Chromosome 6
Chromosome 7
Chromosome 7
Chromosome 8


High hyperdiploid ALL
ETV6/RUNX1-positive ALL
Peak score
Relative peak density per 4 Mb
PTPN1
FOXD4L3
CDC14B
DBC1
Chromosome 9
Chromosome 9
DNTT
Chromosome 10
Chromosome 10
CADM1
Chromosome 11
Chromosome 11
Chromosome 12
Chromosome 12


High hyperdiploid ALL
ETV6/RUNX1-positive ALL
Peak score
Relative peak density per 4 Mb
Chromosome 13
Chromosome 13
Chromosome 14
Chromosome 14
Chromosome 15
Chromosome 15
Chromosome 16
Chromosome 16


High hyperdiploid ALL
ETV6/RUNX1-positive ALL
Peak score
Relative peak density per 4 Mb
Chromosome 17
Chromosome 17
Chromosome 18
Chromosome 18
Chromosome 19
Chromosome 19
Chromosome 20
Chromosome 20

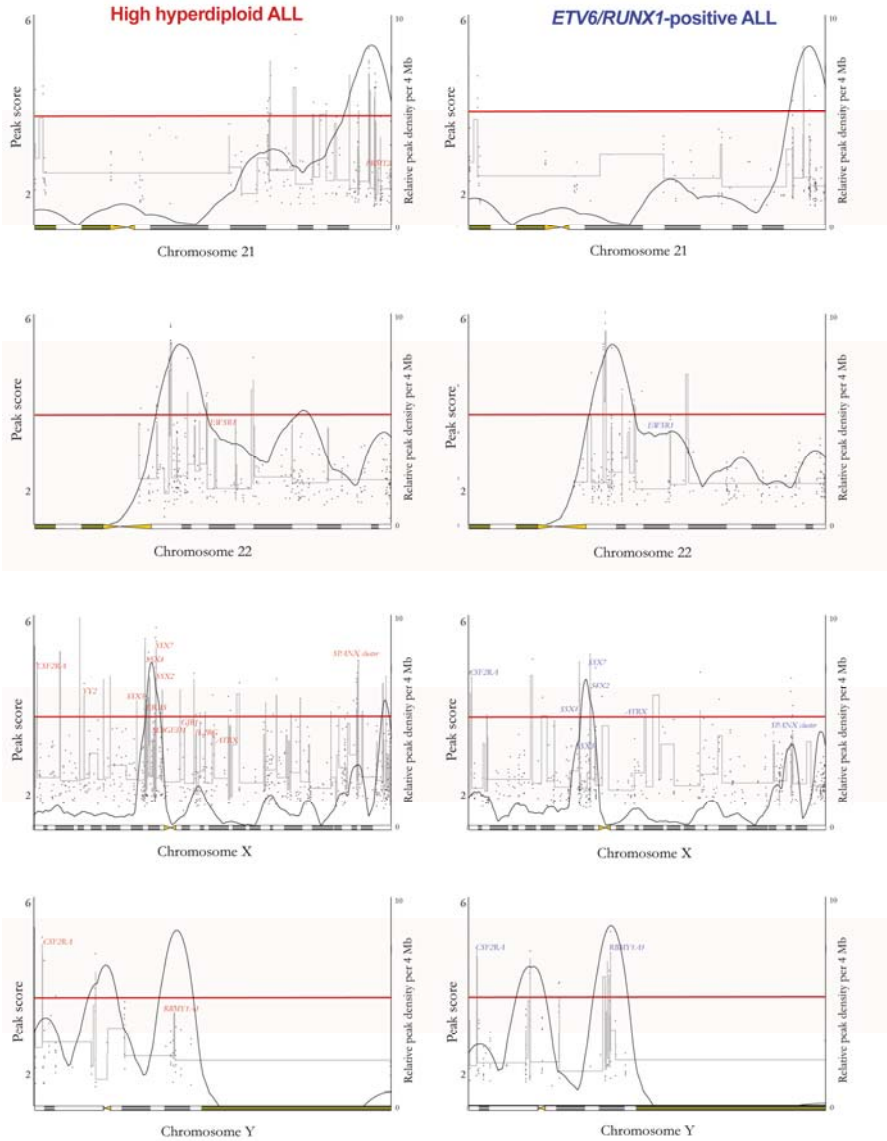
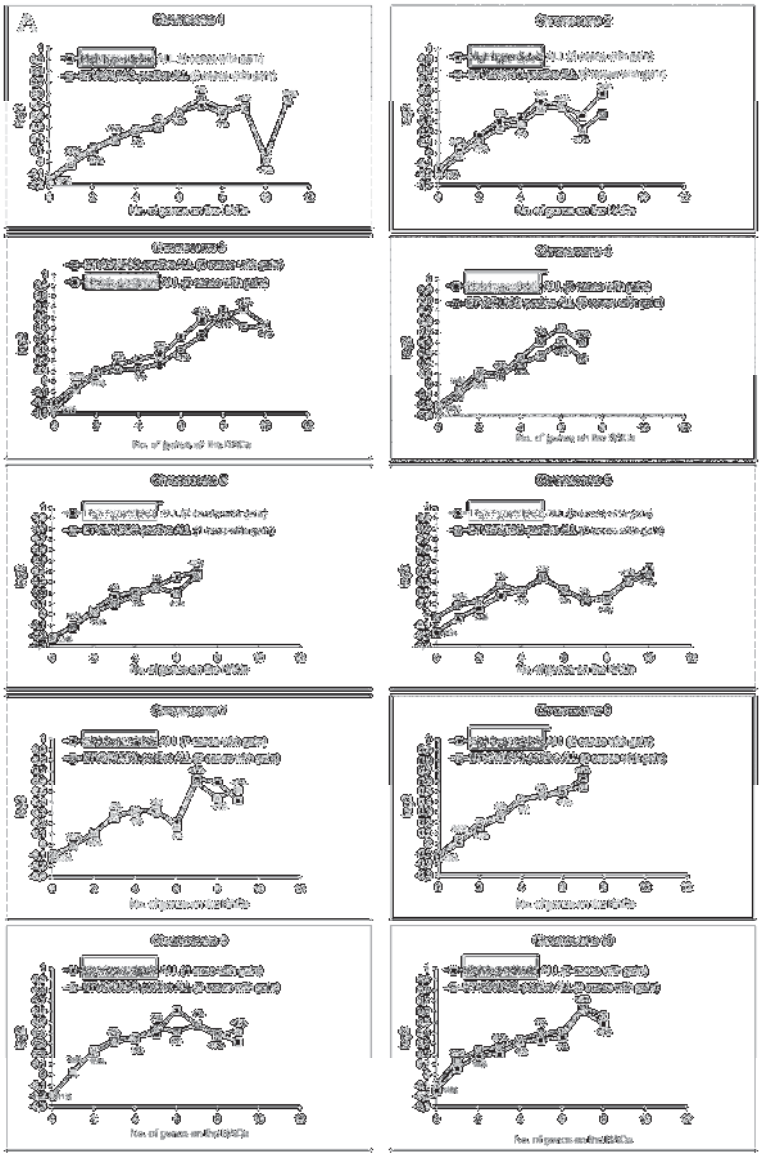
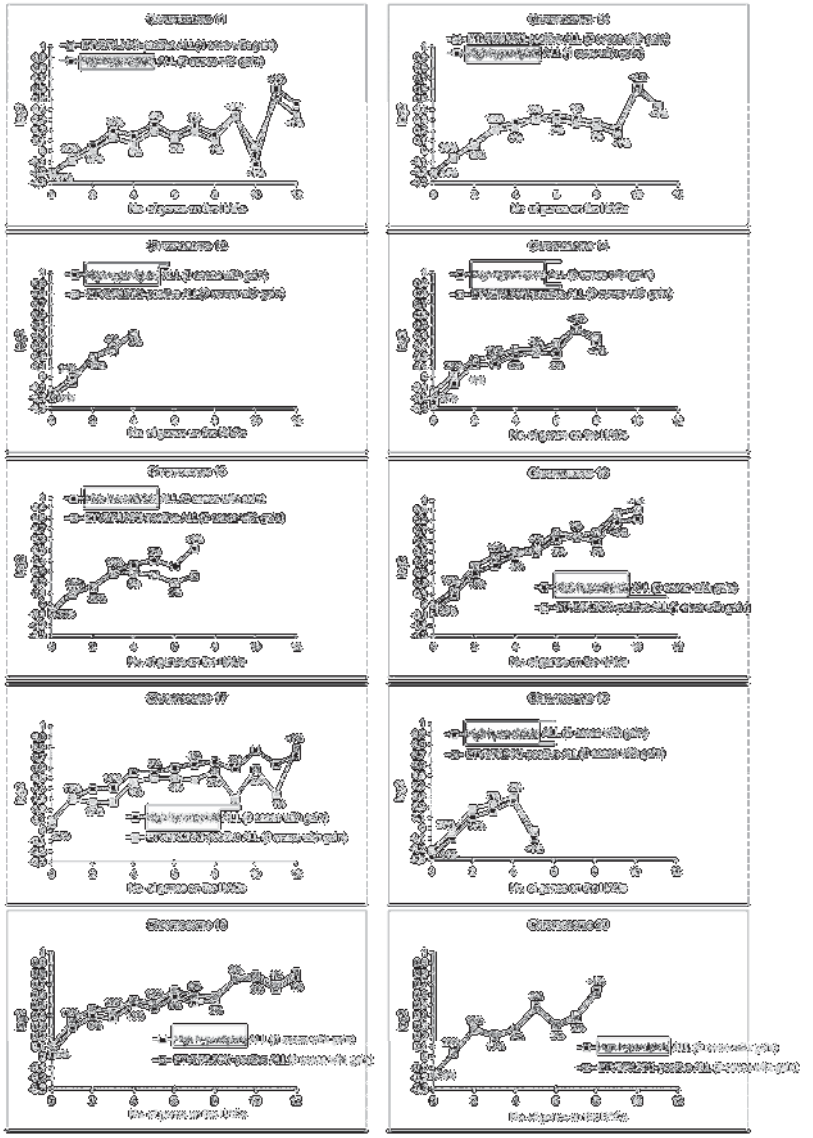
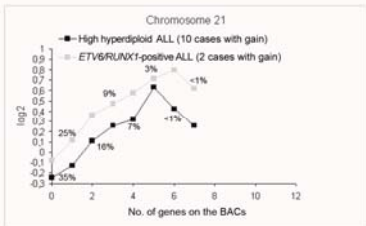
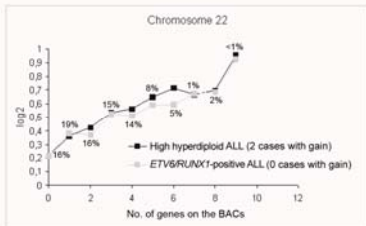


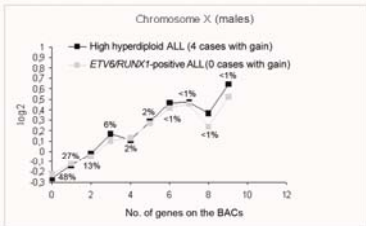
Figure S3. The segmentation analysis identified peaked genes that were highly enriched for methylation. These are labeled in red and blue for high hyperdiploid and *ETV6/RUNX1*-positive ALLs, respectively. The spatial and number distribution of peak-scored probes was plotted using a density estimator with the size determinant set to 4 Mb. In order to designate a chromosomal region as a methylation hotspot at least four peak-scored probes had to be present within the 4 Mb segment analyzed. This cutoff is indicated with the red horizontal line. Note that the *ETV6/RUNX1*-positive ALLs had no peaks for chromosome 8 (Supplementary Material, Table S4).

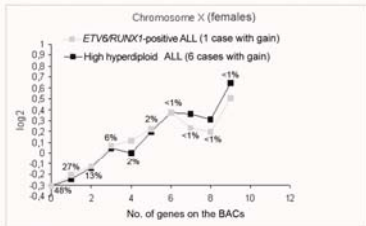


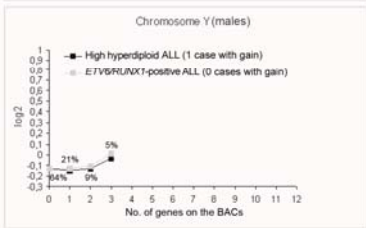


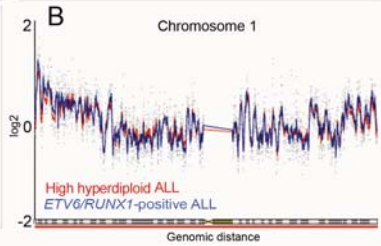

Chromosome 21
High hyperdiploid ALL (10 cases with gain)
ETV6/RUNX1-positive ALL (2 cases with gain)
log2
No. of genes on the BACs

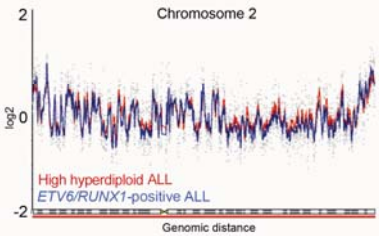

Chromosome 22
High hyperdiploid ALL (2 cases with gain)
ETV6/RUNX1-positive ALL (0 cases with gain)
log2
No. of genes on the BACs

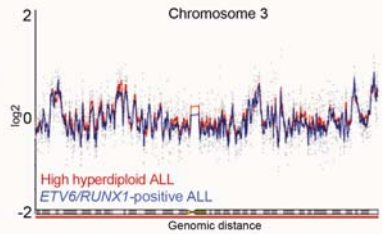

Chromosome X (males)
High hyperdiploid ALL (4 cases with gain)
ETV6/RUNX1-positive ALL (0 cases with gain)
log2
No. of genes on the BACs

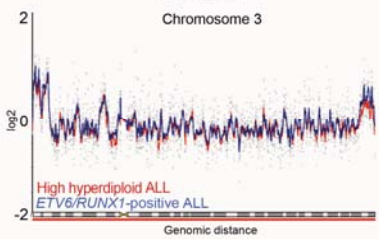

Chromosome X (females)
ETV6/RUNX1-positive ALL (1 case with gain)
High hyperdiploid ALL (6 cases with gain)
log2
No. of genes on the BACs

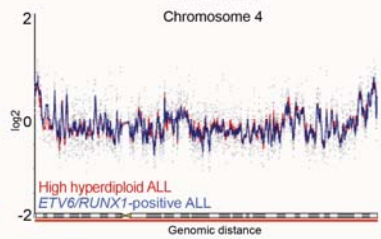

Chromosome Y (males)
High hyperdiploid ALL (1 case with gain)
ETV6/RUNX1-positive ALL (0 cases with gain)
log2
No. of genes on the BACs

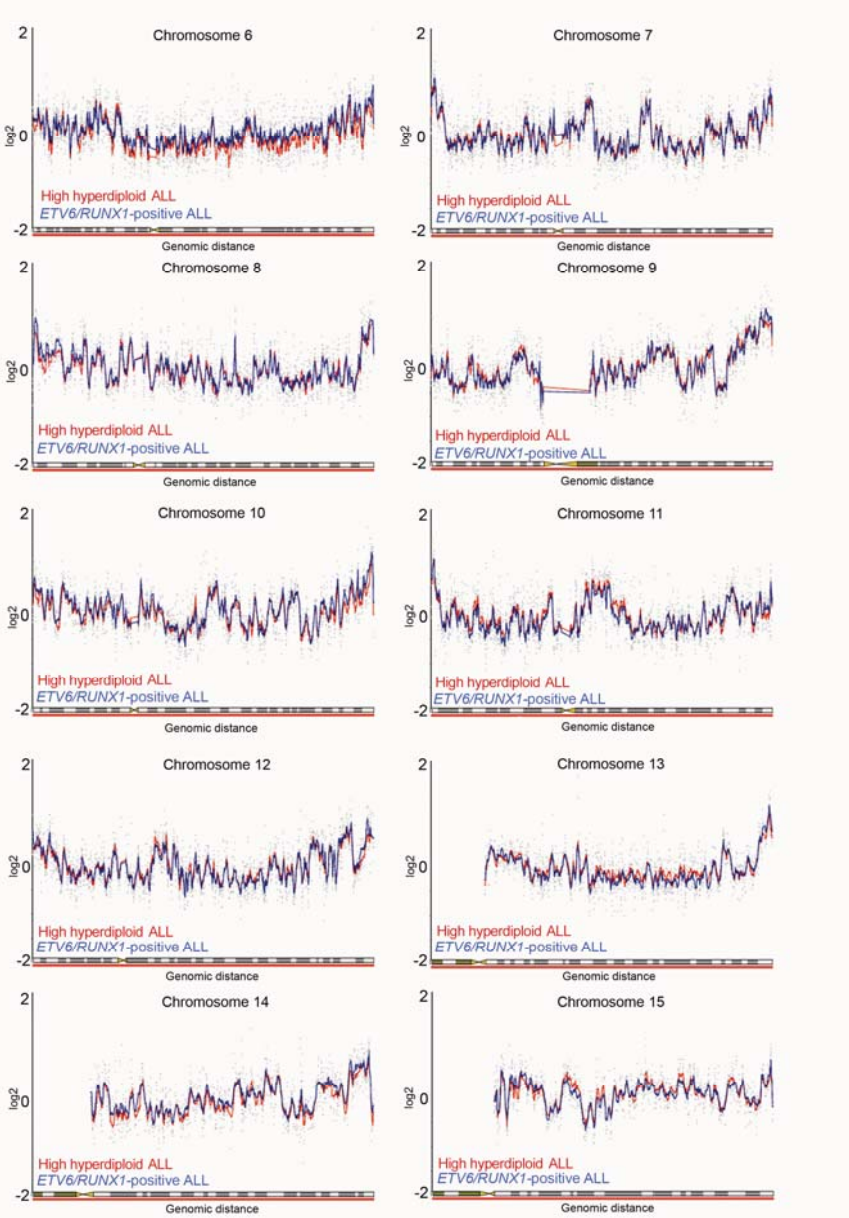

B
Chromosome 1
log2
High hyperdiploid ALL
ETV6/RUNX1-positive ALL
Genomic distance


Chromosome 2
log2
High hyperdiploid ALL
ETV6/RUNX1-positive ALL
Genomic distance


Chromosome 3
log2
High hyperdiploid ALL
ETV6/RUNX1-positive ALL
Genomic distance


Chromosome 3
log2
High hyperdiploid ALL
ETV6/RUNX1-positive ALL
Genomic distance


Chromosome 4
log2
High hyperdiploid ALL
ETV6/RUNX1-positive ALL
Genomic distance


Chromosome 6
Chromosome 7
Chromosome 8
Chromosome 9
Chromosome 10
Chromosome 11
Chromosome 12
Chromosome 13
Chromosome 14
Chromosome 15
log2
2
0
-2
High hyperdiploid ALL
ETV6/RUNX1-positive ALL
Genomic distance

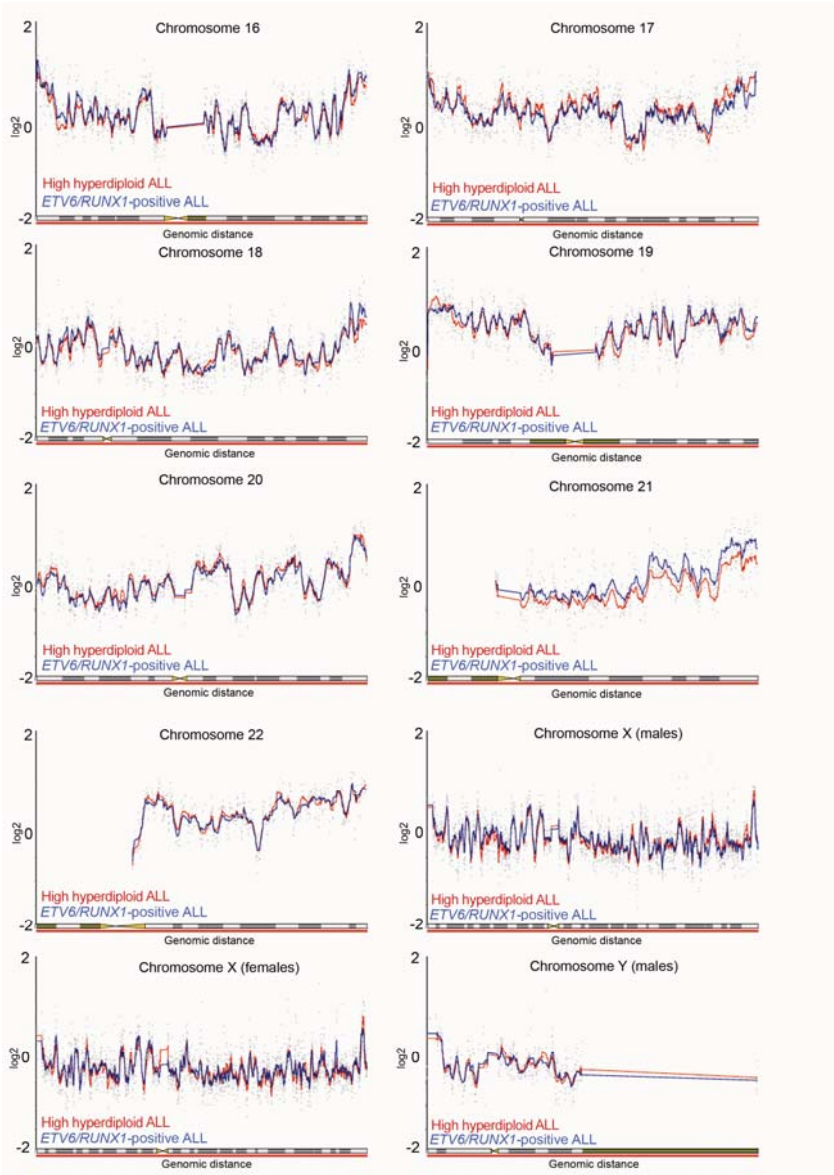


Figure S4. (A) Mean log2 ratios of the *ETV6/RUNX1*-positive and high hyperdiploid ALLs plotted against the number of genes on the bacterial artificial chromosome clones for all chromosomes. The percentages that the clones with different number of genes comprise of the whole chromosome are given above each bacterial artificial chromosome group. **(B)** Mean log2 ratios of the *ETV6/RUNX1*-positive (blue) and high hyperdiploid (red) ALLs plotted against genomic positions on chromosomes 1-22, X and Y

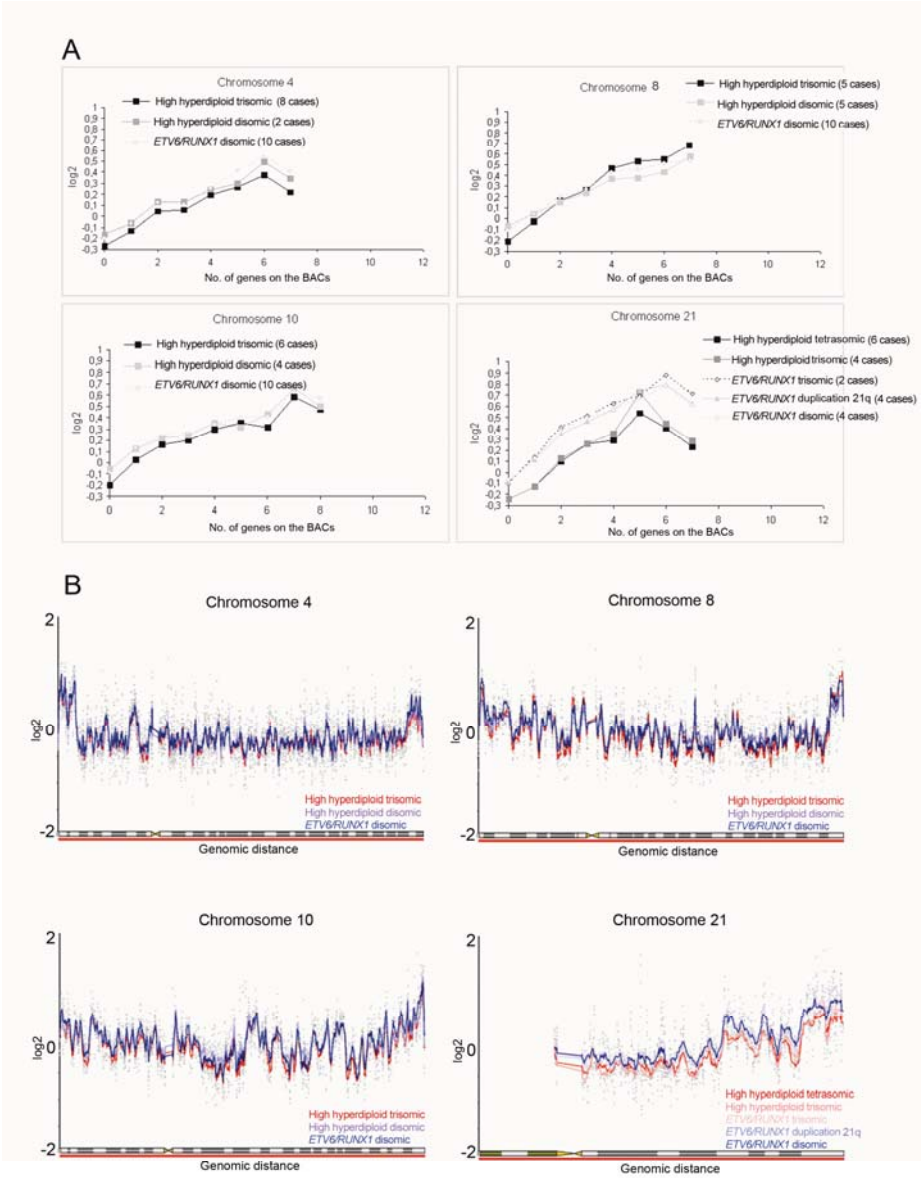


Figure S5. **(A)** Mean log2 ratios of chromosomes 4, 8, 10 and 21 among the different cytogenetic subgroups of the *ETV6/RUNX1*-positive and high hyperdiploid ALLs plotted against the number of genes on the bacterial artificial chromosome clones. **(B)** Mean log2 ratios of chromosomes 4, 8, 10 and 21 plotted against the genomic distance.

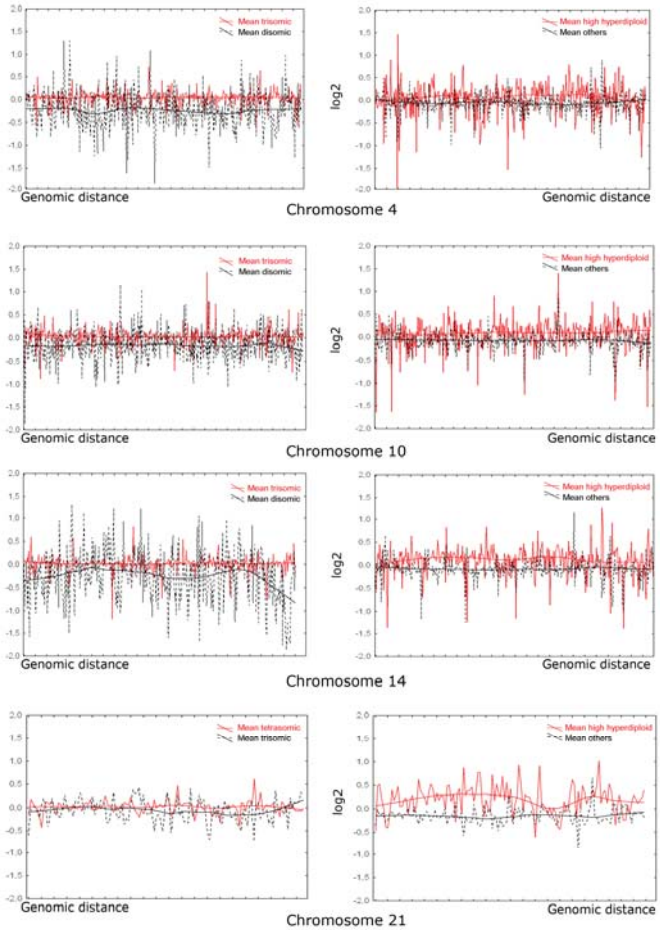
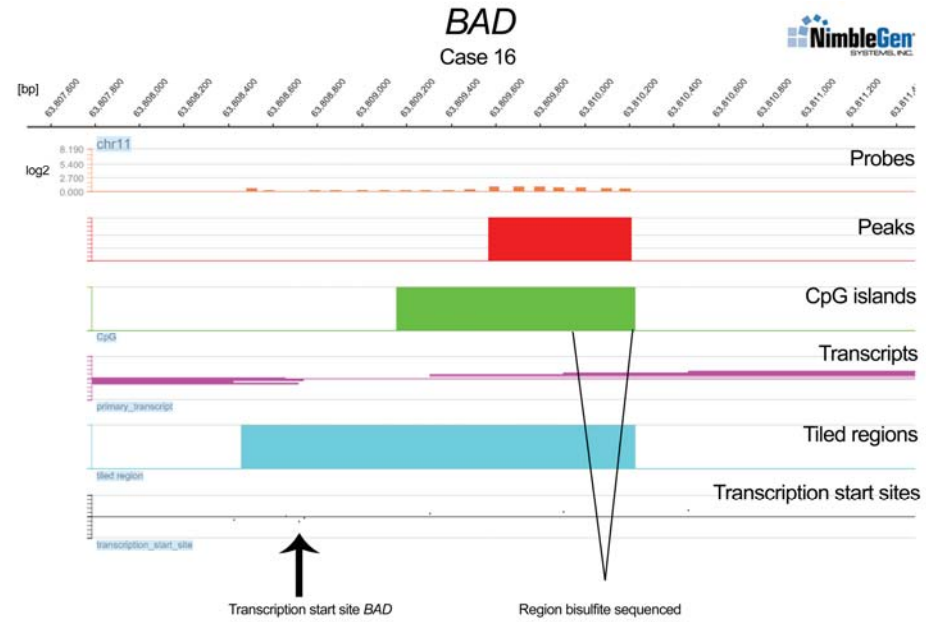
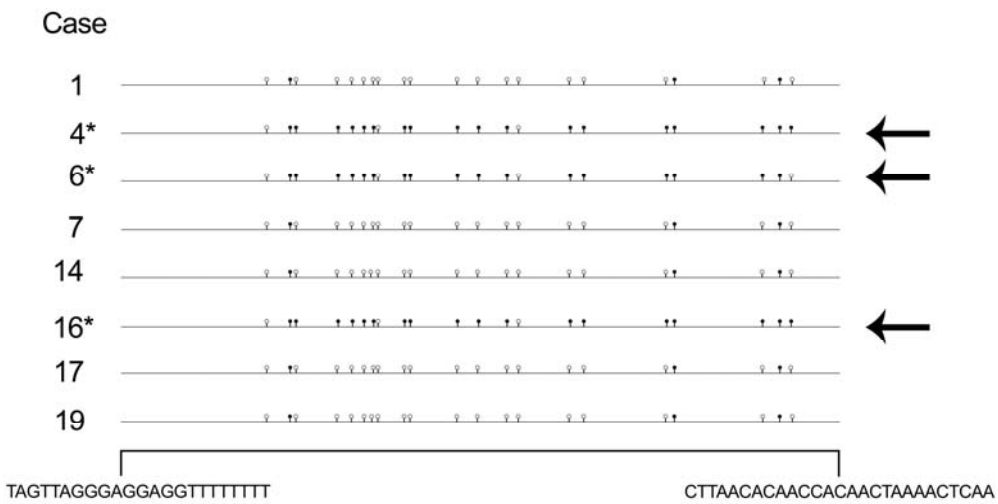


Figure S6. Lowess curves of mean-centered expression ratios of disomies, trisomies and tetrasomies for chromosomes 4, 10, 14 and 21 among the high hyperdiploid cases (left) and between high hyperdiploid cases and all other pediatric B-lineage ALLs (right) analyzed in the study by Andersson *et al.* (1).

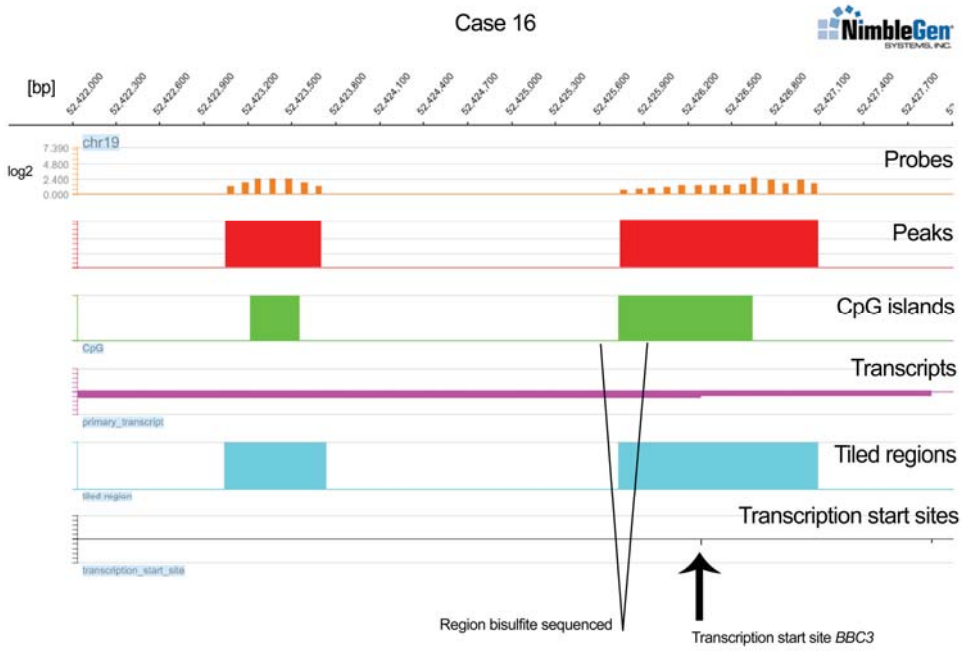
Reference

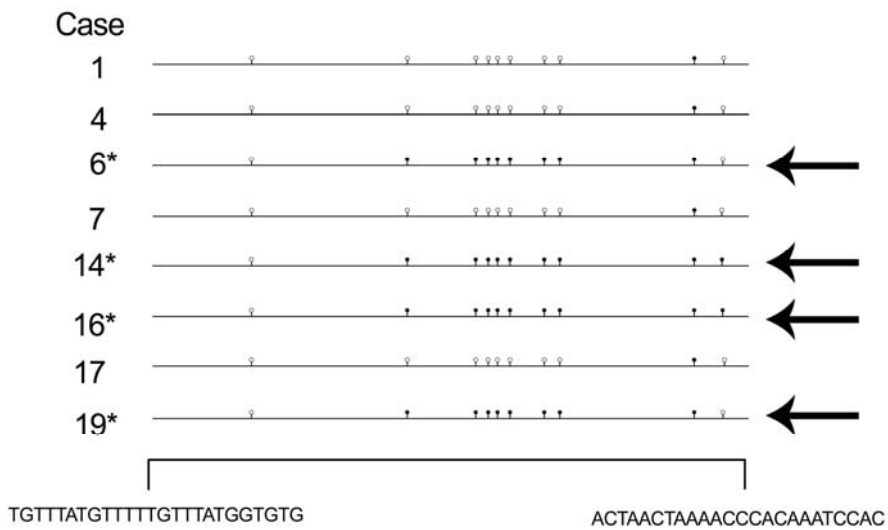
1. Andersson, A., Olofsson, T., Lindgren, D., Nilsson, B., Ritz, C., Edén, P., Lassen, C., Råde, J., Fontes, M., Mörse, H. *et al.* (2005) Molecular signatures in childhood acute leukemia and their correlations to expression patterns in normal hematopoietic subpopulations. *Proc. Natl. Acad. Sci. USA* **102**, 19069-19074.

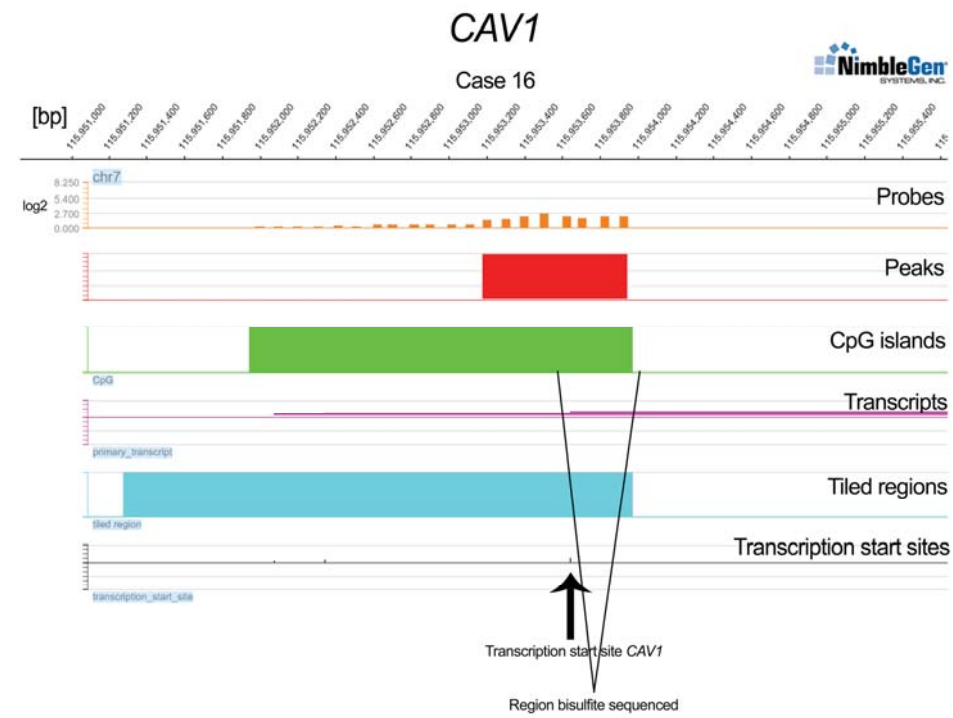

BAD
Case 16
NimbleGen
SYSTEMS, INC.
[bp]
chr11
log2
Probes
Peaks
CpG islands
CpG
Transcripts
primary_transcript
Tiled regions
tiled region
Transcription start sites
transcription_start_site
Transcription start site BAD
Region bisulfite sequenced

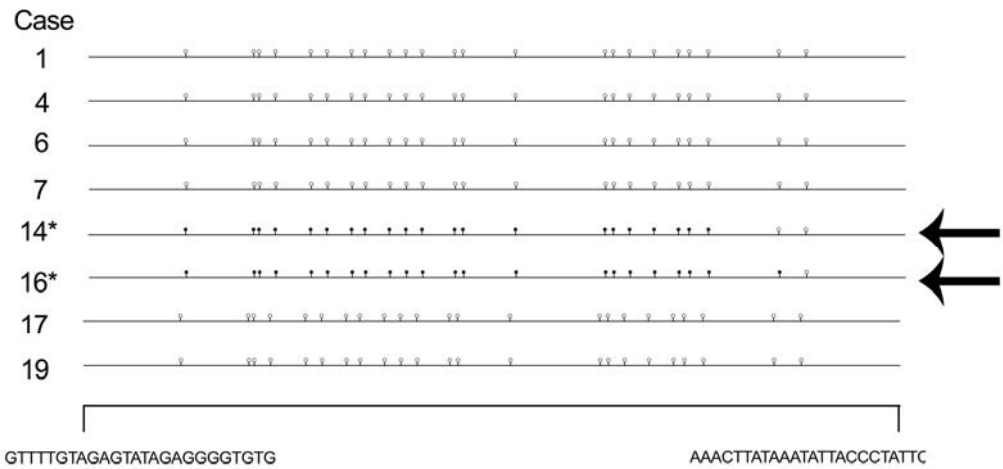

Case
1
4*
6*
7
14
16*
17
19
TAGTTAGGGAGGAGGTTTTTTTT
CTTAACACAACCACAACTAAAACTCAA

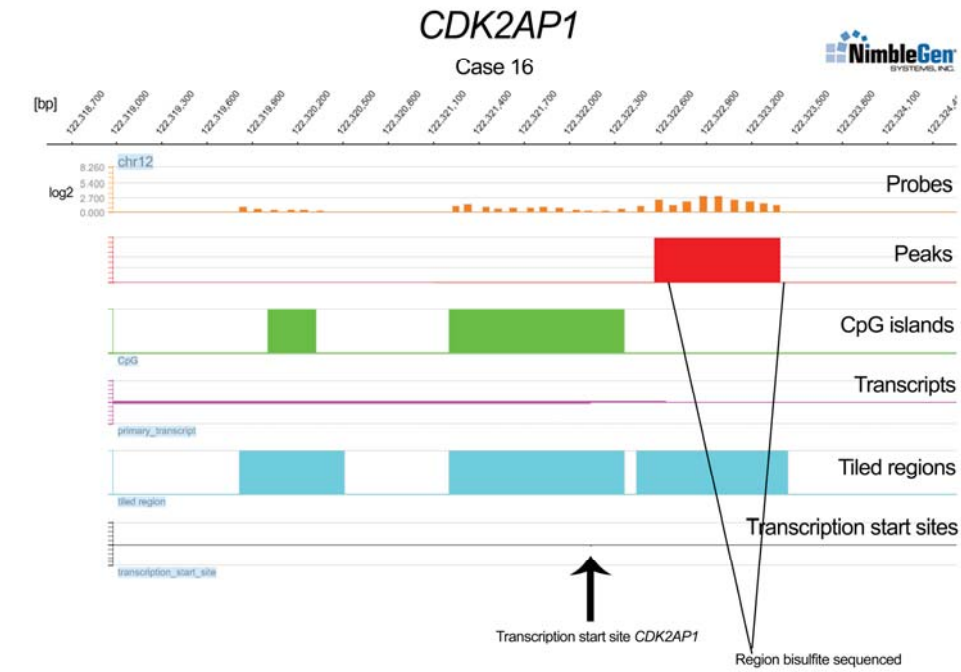
BBC3

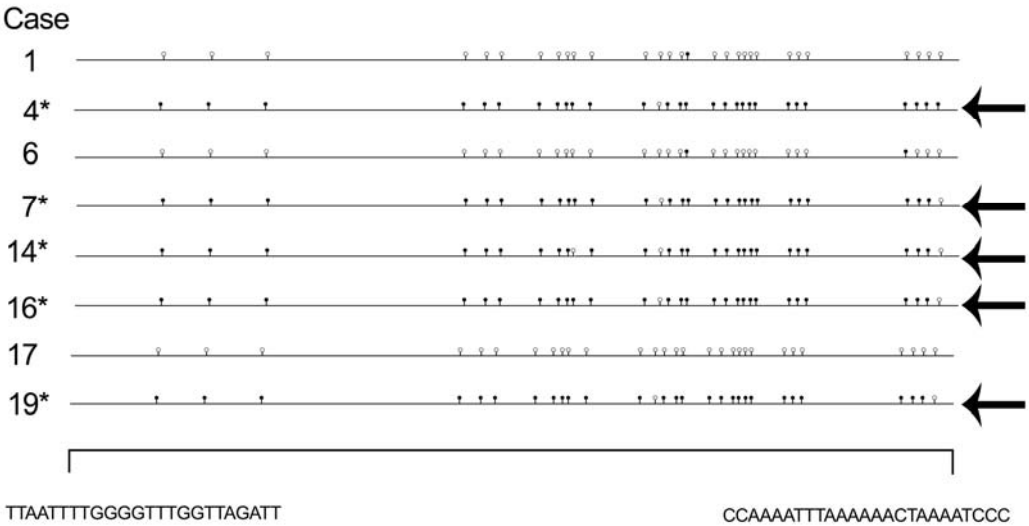

Case 16
NimbleGen
[bp]
chr19
log2
Probes
Peaks
CpG islands
CpG
Transcripts
primary_transcript
Tiled regions
tiled region
Transcription start sites
transcription_start_site
Region bisulfite sequenced
Transcription start site *BBC3*

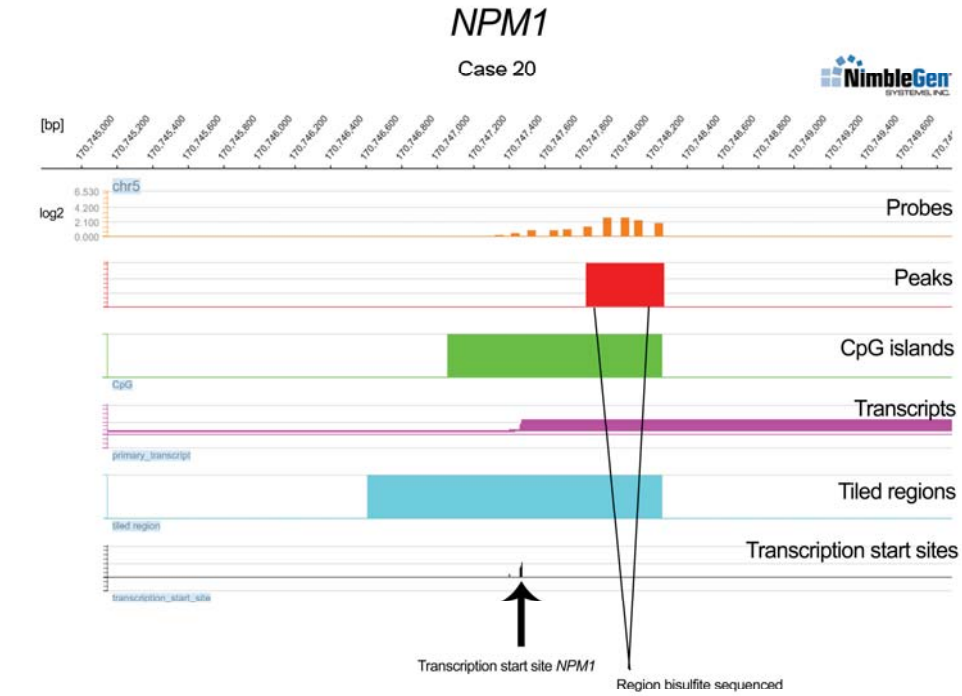

Case
1
4
6*
7
14*
16*
17
19*
TGTTTATGTTTTTGTTTATGGTGTG
ACTAACTAAAACCCACAAATCCAC

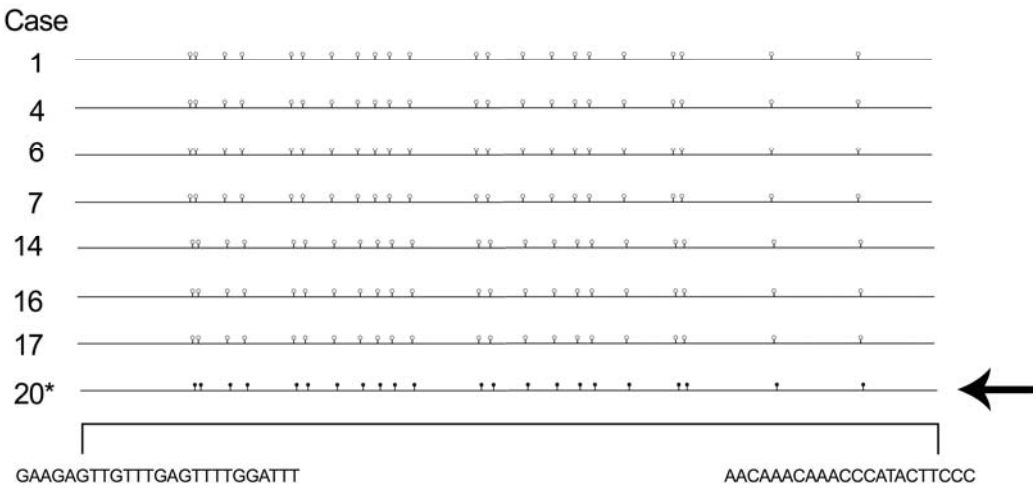

CAV1
Case 16
NimbleGen SYSTEMS INC.
[bp]
chr7
log2
Probes
Peaks
CpG islands
CpG
Transcripts
primary_transcript
Tiled regions
tiled region
Transcription start sites
transcription_start_site
Transcription start site CAV1
Region bisulfite sequenced

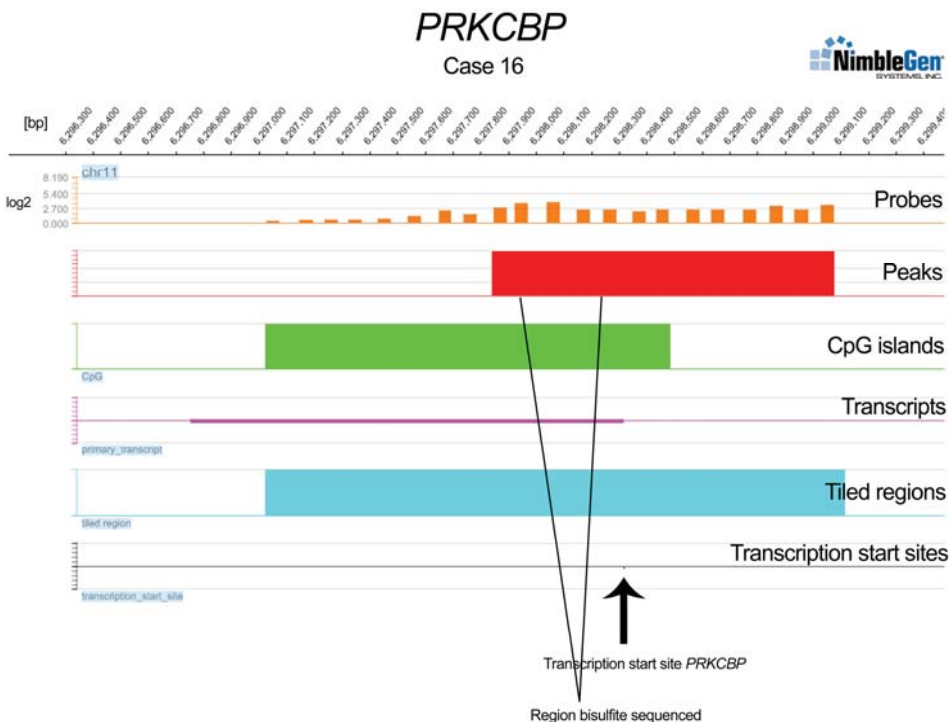

Case
1
4
6
7
14*
16*
17
19
GTTTTGTAGAGTATAGAGGGGTGTG
AAACTTATAAATATTACCCTATTC

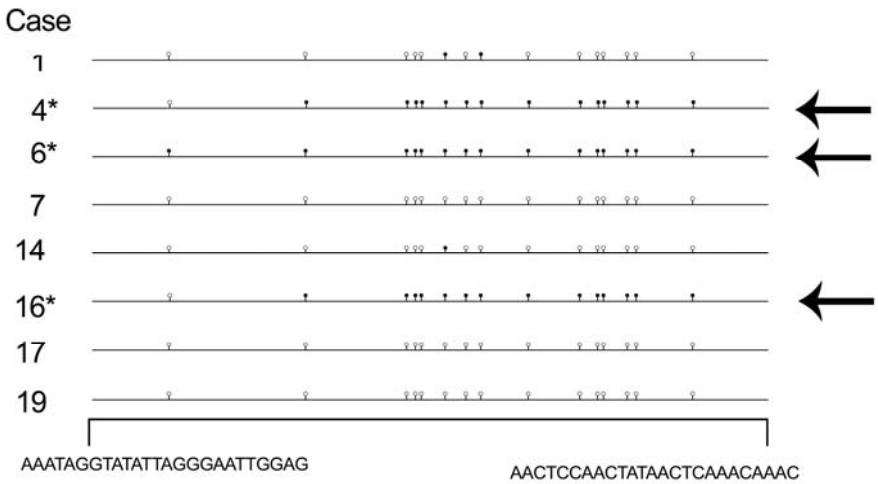

CDK2AP1
Case 16
NimbleGen
SYSTEMS, INC.
[bp]
chr12
log2
Probes
Peaks
CpG islands
CpG
Transcripts
primary_transcript
Tiled regions
tiled region
Transcription start sites
transcription_start_site
Transcription start site CDK2AP1
Region bisulfite sequenced


Case
1
4*
6
7*
14*
16*
17
19*
TTAATTTTGGGGTTTGGTTAGATT
CCAAAATTTAAAAAACTAAAATCCC


NPM1
Case 20
NimbleGen
SYSTEMS INC.
[bp]
chr5
log2
Probes
Peaks
CpG islands
CpG
Transcripts
primary_transcript
Tiled regions
tiled region
Transcription start sites
transcription_start_site
Transcription start site NPM1
Region bisulfite sequenced


Case
1
4
6
7
14
16
17
20*
GAAGAGTTGTTTGAGTTTTGGATTT
AACAAACAAACCCATACTTCCC


PRKCBP
Case 16
NimbleGen
[bp]
log2
chr11
Probes
Peaks
CpG islands
CpG
Transcripts
primary_transcript
Tiled regions
tiled region
Transcription start sites
transcription_start_site
Transcription start site PRKCBP
Region bisulfite sequenced


Case
1
4*
6*
7
14
16*
17
19
AAATAGGTATATTAGGGAATTGGAG
AACTCCAACTATAACTCAAACAAAC

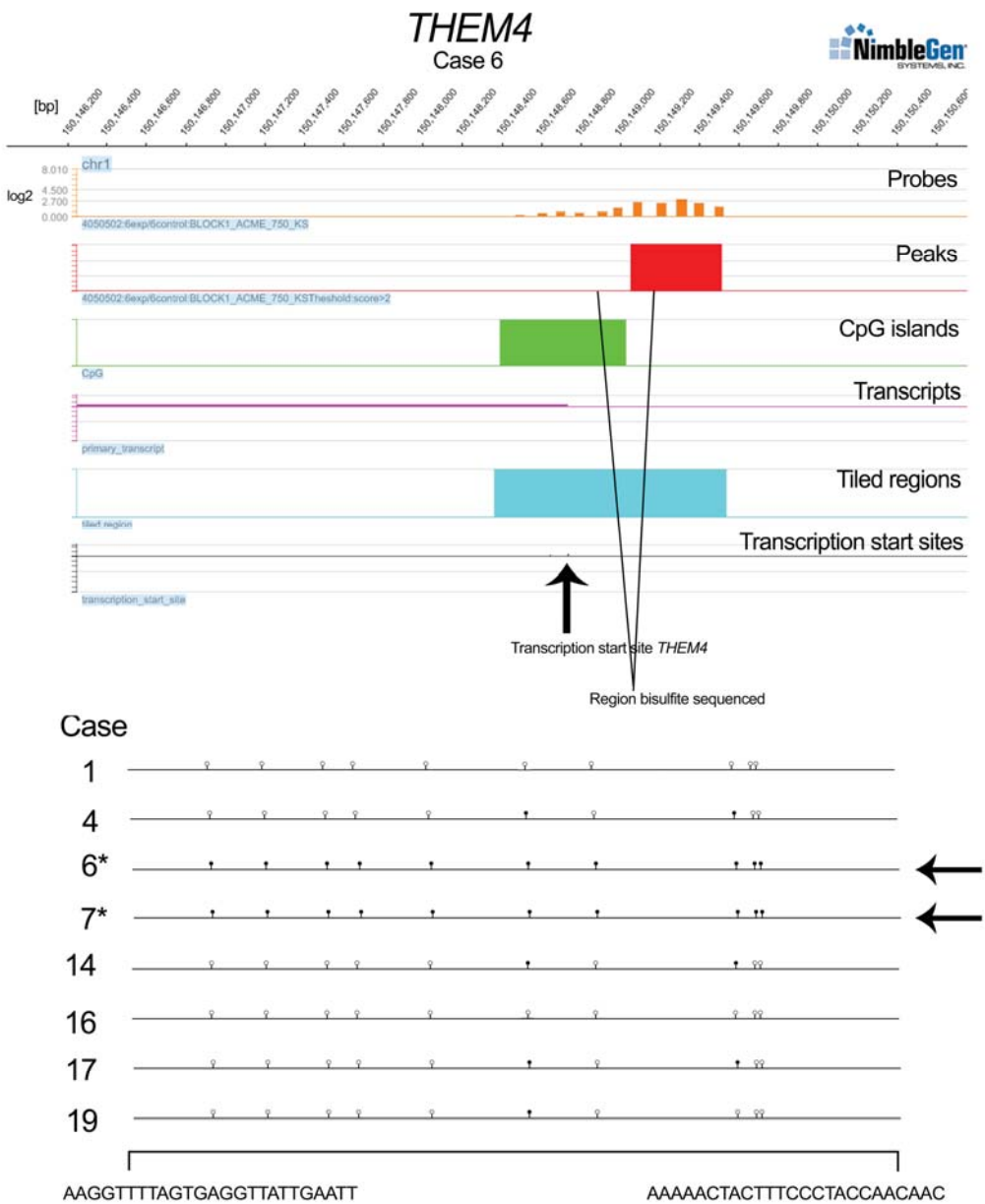


Figure S7. Bisulfite sequencing validating the array results. Cases harboring a methylation peak detected by the promoter array are marked with an asterisk. Sequences demonstrated to be hypermethylated by bisulfite sequencing are indicated with arrows. Primer sequences used are given below each gene.

Table S1. Genes that previously have been reported to be frequently hypermethylated in malignant disorders[a]

Gene	Function[b]	Location[b]	Tissue/organ[a]
ABCB1	Drug resistance	7q21.1	Hematopoietic
ABCB4	Drug resistance	7q21.1	Various
APC	Cell cycle	5q21-22	Various
AR	Cell signaling	Xq11.2-12	Prostate
BRCA1	DNA repair	17q21	Breast, ovary
CALCA	Cell signaling	11p15.1-15.2	Various
CASP8	Apoptosis	2q33-34	Neuronal
CAV1	Transport	7q31.1	Breast
CCNA1	Cell cycle	13q12.3-13	Various
CD44	Cell adhesion	11p13	Prostate
CDH1	Cell adhesion	16q22.1	Various
CDH13	Cell adhesion	16q24.2-24.3	Breast, lung
CDKN1B	Cell cycle	12p12-13.1	Various
CDKN1C	Cell cycle	11p15.5	Stomach
CDKN2A	Cell cycle	9p21	Various
CDKN2B	Cell cycle	9p21	Hematopoietic
CFTR	Ion transport	7q31.2	Various
DAPK1	Apoptosis	9q34.1	Various
DKK1	Cell signaling	10q11.2	Colon
EDNRB	Cell signaling	13q22	Prostate
EPHA3	Cell signaling	3p11.2	Hematopoietic
EPO	Cell signaling	7q22	Colon
ESR1	Cell signaling	6q25.1	Breast
EXT1	Cell cycle	8q24.1	Hematopoietic, skin
FABP3	Cell proliferation	1p32-33	Breast, colon
FAT	Cell adhesion	4q35	Colon
FHIT	Cell cycle	3p14.2	Esophagus
GATA4	Transcription regulation	8p22-23.1	Colon, stomach
GATA5	Transcription regulation	20q13.3	Colon, stomach
GJB2	Cell signaling	13q11-12	Breast
GPC3	Morphogenesis	Xq26.1	Ovary
GSTP1	Transferase activity	11q13	Various

H19	Unknown	11p15.5	Kidney
HIC1	Transcription regulation	17p13.3	Various
HOXA9	Transcription regulation	7p14-15	Neuronal
ID4	Transcription regulation	6p21-22	Hematopoietic, stomach
IGF2	Cell proliferation	11p15.5	Colon, hematopoietic
IGFBP3	Apoptosis	7p12-13	Lung, skin
IGFBP7	Cell proliferation	4q12	Liver
LMNA	Morphogenesis	1q21.2-21.3	Hematopoietic
LTB4R	Cell signaling	14q11.2-12	Various
MGMT	DNA repair	10q26	Various
MLH1	DNA repair	3p21.3	Various
MME	Cell signaling	3q25.1-25.2	Prostate
MT1A	Ion binding	16q13	Liver
MUC2	Extracellular structure	11p15.5	Colon
MYOD1	Cell differentiation	11p15.4	Various
NEFL	Morphogenesis	8p21	Brain
PAX6	Transcription regulation	11p13	Various
PGR	Cell signaling	11q22-23	Breast
PLAU	Cell signaling	10q24	Breast
PRDM2	Transcription regulation	1p36.2	Various
PRKCDBP	Cell cycle	11p15.4	Breast, lung
PRLR	Cell signaling	5p13-14	Breast
PTGS2	Cell motility	1q25.2-25.3	Colon, stomach
PYCARD	Apoptosis	16p11.2-12	Breast
P14ARF	Cell cycle	9p21	Various
RARB	Transcription regulation	3p24	Various
RASSF1	Cell cycle	3p21.3	Various
RASSF5	Apoptosis	1q32.1	Lung
RB1	Cell cycle	13q14.2	Various
RBP1	Transport	3q23	Various
SFN	Apoptosis	1p36.1	Breast, colon
SFRP1	Cell signaling	8p11.1-12	Colon
SLC5A5	Ion transport	19p12-13.2	Thyroid
SLC5A8	Ion transport	12q23.2	Brain, colon
SOCS1	Cell signaling	16p13.1	Hematopoietic, liver
SOCS3	Cell signaling	17q25.3	Lung

STK11	Cell cycle	19p13.3	Various
SYK	Cell signaling	9q22	Breast
TERT	Telomere maintenance	15p15.3	Various
TES	Ion binding	7q31.2	Hematopoietic
THBS1	Cell adhesion	15q15	Brain
TIMP3	Apoptosis	22q12.1-13.2	Various
TMEFF2	Unknown	2q32.3	Bladder, colon
TP73	Apoptosis	1p36.3	Hematopoietic
TSHR	Cell signaling	14q31	Thyroid
TUSC3	Glycosylation	8p22	Various
VCAN	Cell adhesion	5q14.3	Colon
VHL	Apoptosis	3p25-26	Various
WIF1	Cell signaling	12q14.3	Colon, lung
WRN	DNA repair	8p11.2-12	Various
WT1	Transcription regulation	11p13	Various

[a]Retrieved from the Genes Methylated in Cancer Website at MD Anderson Cancer Center, The University of Texas (www.mdanderson.org/departments/methylation/) and from a review by Esteller (1).

[b]Function and location are based on data in NCBI Entrez Gene (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene). Only one major function per gene is given.

Reference

1. Esteller, M. (2007) Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat. Rev. Genet.*, **8**, 286-298.

Table S2. Hypermethylated genes in B-lineage ALL[a]

Gene	Function[b]	Location[c]
ABCB1	Drug resistance	7q21.1
ABL1	Cell signaling	9q34.1
ADAMTS1	Proteolysis	21q21.2
ADAMTS5	Proteolysis	21q21.3
AHR	Apoptosis	7p15
APAF1	Apoptosis	12q23
APC	Cell cycle	5q21-22
CADM1	Apoptosis	11q23.2
CALCA	Cell signaling	11p15.1-15.2
CCND2	Cell cycle	12p13
CD44	Cell adhesion	11p13
CDC2	Cell cycle	10q21.1
CDH1	Cell adhesion	16q22.1
CDH13	Cell adhesion	16q24.2-24.3
CDKN1A	Cell cycle	6p21.2
CDKN1C	Cell cycle	11p15.5
CDKN2A	Cell cycle	9p21
CDKN2B	Cell cycle	9p21
CHFR	Cell cycle	12q24.33
DACT1	Cell signaling	14q23.1
DAPK1	Apoptosis	9q34.1
DBC1	Apoptosis	9q32-33
DCC	Apoptosis	18q21.3
DDX51	RNA processing	12q24.3
DIABLO	Apoptosis	12q24.3
DKK3	Cell signaling	11p15.2
DLC1	Cell adhesion	8p22
EGR4	Transcription regulation	2p13
ESR1	Cell signaling	6q25.1
EXT1	Cell cycle	8q24.1
FHIT	Cell cycle	3p14.2
GATA5	Transcription regulation	20q13.3
GGH	Peptidase	8q12.3
GSTP1	Transferase activity	11q13

HCK	Morphogenesis	20q11-12
HIC1	Transcription regulation	17p13.3
KCNK2	Ion transport	1q41
KLK10	Cell cycle	19q13.3-13.4
LATS1	Cell cycle	6q24-25.1
LATS2	Cell cycle	13q11-12
LMNA	Morphogenesis	1q21.2-21.3
LRP1B	Proteolysis	2q21.2
MGMT	DNA repair	10q26
MME	Cell signaling	3q25.1-25.2
MOS	Cell cycle	8q11
MSH6	Apoptosis	2p16
NKX6-1	Morphogenesis	4q21.2-22
NNAT	Morphogenesis	20q11.2-12
NOPE	Unknown	15q22.3
P14ARF	Cell cycle	9p21
PACRG	Unknown	6q26
PARK2	Proteolysis	6q25.2-27
PAX6	Transcription regulation	11p13
PCDHGA12	Cell adhesion	5q31
PGR	Cell signaling	11q22-23
PPP1R13B	Apoptosis	14q32.3
PRDM2	Transcription regulation	1p36.2
PTEN	Cell cycle	10q23.3
PTPN6	Cell signaling	12p13
PYCARD	Apoptosis	16p11.2-12
RARB	Transcription regulation	3p24
RPRM	Cell cycle	2q23.3
RUNDC3B	Unknown	7q21.1
SEPT4	Apoptosis	17q22-23
SFRP1	Cell signaling	8p11.1-12
SFRP2	Cell signaling	4q31.3
SFRP4	Cell signaling	7p14.1
SFRP5	Cell signaling	10q24.1
SLC2A14	Transport	12p13.3
SLC19A1	Transport	21q22.3

SMC1A	Cell cycle	Xp11.2
SMC1B	Cell cycle	22q13.3
SYK	Cell signaling	9q22
THBS1	Cell adhesion	15q15
TMP3	Apoptosis	22q12.1-13.2
TP53	Apoptosis	17p13.1
TP73	Apoptosis	1p36.3
TUSC3	Glycosylation	8p22
WIF1	Cell signaling	12q14.3
WNT5A	Cell signaling	3p14-21
WT1	Transcription regulation	11p13

[a]The genes were identified through a PubMed search using the terms "pediatric", "ALL", "CpG" and "methylation". Initially, more than 56 articles were obtained, but when including only those that reported hypermethylated genes in B-lineage ALL only 40 remained (1-40) Based on these publications, a total of 81 genes have been shown to be hypermethylated in B-lineage ALL.

[b]Function is based on data produced by the EASE software (41) or found in NCBI Entrez Gene (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene). Only one major function per gene is given.

[c]Location is based on information in NCBI Entrez Gene (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene).

References

1. Paulsson, K., An, Q., Moorman, A.V., Parker, H., Molloy, G., Davies, T., Griffiths, M., Ross, F.M., Irving, J., Harrison, C.J. *et al.* (2009) Methylation of tumour suppressor gene promoters in the presence and absence of transcriptional silencing in high hyperdiploid acute lymphoblastic leukaemia. *Br. J. Haematol.*, **144**, 838-847.

2. Taylor, K.H., Kramer, R.S., Davis, J.W., Guo, J., Duff, D.J., Xu, D., Caldwell, C.W. and Shi, H. (2007) Ultradeep bisulfite sequencing analysis of DNA methylation patterns in multiple gene promoters by 454 sequencing. *Cancer Res.*, **67**, 8511-8518.

3. Taylor, K.H., Pena-Hernandez, K.E., Davis, J.W., Arthur, G.L., Duff, D.J., Shi, H., Rahmatpanah, F.B., Sjahputera, O. and Caldwell, C.W. (2007) Large-scale CpG methylation analysis identifies novel candidate genes and reveals methylation hotspots in acute lymphoblastic leukemia. *Cancer Res.*, **67**, 2617-2625.

4. Hoshino, K., Quintás-Cardama, A., Yang, H., Sanchez-Gonzalez, B. and Garcia-Manero, G. (2007) Aberrant DNA methylation of the Src kinase Hck, but not of Lyn, in Philadelphia chromosome negative acute lymphocytic leukemia. *Leukemia*, **21**, 906-911.
5. Roman-Gomez, J., Jimenez-Velasco, A., Cordeu, L., Vilas-Zornoza, A., San Jose-Eneriz, E., Garate, L., Castillejo, J.A., Martin, V., Prosper, F., Heiniger, A. *et al.* (2007) *WNT5A*, a putative tumour suppressor of lymphoid malignancies, is inactivated by aberrant methylation in acute lymphoblastic leukaemia. *Eur. J. Cancer*, **43**, 2736-2746.
6. Cheng, Q., Cheng, C., Crews, K.R., Ribeiro, R.C., Pui, C.H., Relling, M.V. and Evans, W.E. (2006) Epigenetic regulation of human γ-glutamyl hydrolase activity in acute lymphoblastic leukemia cells. *Am. J. Hum. Genet.*, **79**, 264-274.
7. Roman-Gomez, J., Jimenez-Velasco, A., Agirre, X., Castillejo, J.A., Navarro, G., Calasanz, M.J., Garate, L., San Jose-Eneriz, E., Cordeu, L., Prosper, F. *et al.* (2006) CpG island methylator phenotype redefines the prognostic effect of t(12;21) in childhood acute lymphoblastic leukemia. *Clin. Cancer Res.*, **12**, 4845-4850.
8. Pãixao, V.A., Vidal, D.O., Caballero, O.L., Vettore, A.L., Tone, L.G., Ribeiro, K.B. and Lopes, L.F. (2006) Hypermethylation of CpG island in the promoter region of *CALCA* in acute lymphoblastic leukemia with central nervous system (CNS) infiltration correlates with poorer prognosis. *Leuk. Res.*, **30**, 891-894.
9. Mulero-Navarro, S., Carvajal-Gonzalez, J.M., Herranz, M., Ballestar, E., Fraga, M.F., Ropero, S., Esteller, M. and Fernandez-Salguero, P.M. (2006) The dioxin receptor is silenced by promoter hypermethylation in human acute lymphoblastic leukemia through inhibition of Sp1 binding. *Carcinogenesis,* **27,** 1099-1104.
10. Stam, R.W., den Boer, M.L., Passier, M.M., Janka-Schaub, G.E., Sallan, S.E., Armstrong, S.A. and Pieters, R. (2006) Silencing of the tumor suppressor gene *FHIT* is highly characteristic for *MLL* gene rearranged infant acute lymphoblastic leukemia. *Leukemia,* **20,** 264-271.
11. Agirre, X., Román-Gómez, J., Vázquez, I., Jiménez-Velasco, A., Garate, L., Montiel-Duarte, C., Artieda, P., Cordeu, L., Lahortiga, I., Calasanz, M.J. *et al.* (2006) Abnormal methylation of the common *PARK2* and *PACRG* promoter is associated with downregulation of gene expression in acute lymphoblastic leukemia and chronic myeloid leukemia. *Int. J. Cancer*, **118**, 1945-1953.
12. Gutiérrez, M.I., Siraj, A.K., Ibrahim, M.M., Hussain, A. and Bhatia, K. (2005) Childhood and adult ALL: Differences in epigenetic lesions associated with cell cycle genes. *Am. J. Hematol.*, **80**, 158-160.

13. Yang, Y., Takeuchi, S., Hofmann, W.K., Ikezoe, T., van Dongen, J.J., Szczepanski, T., Bartram, C.R., Yoshino, N., Taguchi, H. and Koeffler, H.P. (2006) Aberrant methylation in promoter-associated CpG islands of multiple genes in acute lymphoblastic leukemia. *Leuk. Res*., **30,** 98-102.

14. Canalli, A.A., Yang, H., Jeha, S., Hoshino, K., Sanchez-Gonzalez, B., Brandt, M., Pierce, S., Kantarjian, H., Issa, J.P. and Garcia-Manero, G. (2005) Aberrant DNA methylation of a cell cycle regulatory pathway composed of P73, P15 and P57KIP2 is a rare event in children with acute lymphocytic leukemia. *Leuk. Res*., **29,** 881-885.

15. Agrelo, R., Setien, F., Espada, J., Artiga, M.J., Rodriguez, M., Perez-Rosado, A., Sanchez-Aguilera, A., Fraga, M.F., Piris, M.A. and Esteller, M. (2005) Inactivation of the *Lamin A/C* gene by CpG island promoter hypermethylation in hematologic malignancies, and its association with poor survival in nodal diffuse large B-cell lymphoma. *J. Clin.* Oncol., **23,** 3940-3947.

16. Matsuno, N., Hoshino, K., Nanri, T., Kawakita, T., Suzushima, H., Kawano, F., Mitsuya, H. and Asou, N. (2005) *p15* mRNA expression detected by real-time quantitative reverse transcriptase-polymerase chain reaction correlates with the methylation density of the gene in adult acute leukemia. *Leuk. Res*., **29,** 557-564.

17. Tsellou, E., Troungos, C., Moschovi, M., Athanasiadou-Piperopoulou, F., Polychronopoulou, S., Kosmidis, H., Kalmanti, M., Hatzakis, A., Dessypris, N., Kalofoutis, A. *et al.* (2005) Hypermethylation of CpG islands in the promoter region of the *p15INK4B* gene in childhood acute leukaemia. *Eur. J. Cancer*, **41,** 584-589.

18. Sahu, G.R. and Das, B.R. (2005) Alteration of p73 in pediatric de novo acute lymphoblastic leukemia. *Biochem. Biophys. Res. Commun*., **327**, 750-755.

19. Zemliakova, V.V., Strelnikov, V.V., Zborovskaia, I.B., Balukova, O.V., Maiorova, O.A., Vasilev, E.V., Zaletaev, D.V. and Nemtsova, M.V. (2004) [Abnormal methylation of p16/CDKN2A AND p14/ARF genes GpG islands in non-small cell lung cancer and in acute lymphoblastic leukemia]. *Mol. Biol.* (Mosc), **38**, 966-972.

20. Scholz, C., Nimmrich, I., Burger, M., Becker, E., Dorken, B., Ludwig, W.D. and Maier, S. (2005) Distinction of acute lymphoblastic leukemia from acute myeloid leukemia through microarray-based DNA methylation analysis. *Ann. Hematol.*, **84,** 236-244.

21. Ropero, S., Setien, F., Espada, J., Fraga, M.F., Herranz, M., Asp, J., Benassi, M.S., Franchi, A., Patino, A., Ward, L.S. *et al.* (2004) Epigenetic loss of the familial tumor-suppressor gene exostosin-1 (EXT1) disrupts heparan sulfate synthesis in cancer cells. *Hum. Mol. Genet*., **13,** 2753-2765.

22. Roman-Gomez, J., Jimenez-Velasco, A., Agirre, X., Castillejo, J.A., Navarro, G., Barrios, M., Andreu, E.J., Prosper, F., Heiniger, A. and Torres, A. (2004) Transcriptional silencing of the *Dickkopfs-3* (*Dkk-3*) gene by CpG hypermethylation in acute lymphoblastic leukaemia. *Br. J. Cancer,* **91,** 707-713.

23. Matsushita, C., Yang, Y., Takeuchi, S., Matsushita, M., Van Dongen, J.J., Szczepanski, T., Bartram, C.R., Seo, H., Koeffler, H.P. and Taguchi, H. (2004) Aberrant methylation in promoter-associated CpG islands of multiple genes in relapsed childhood acute lymphoblastic leukemia. *Oncol. Rep.*, **12,** 97-99.

24. Zheng, S., Ma, X., Zhang, L., Gunn, L., Smith, M.T., Wiemels, J.L., Leung, K., Buffler, P.A. and Wiencke, J.K. (2004) Hypermethylation of the 5' CpG island of the *FHIT* gene is associated with hyperdiploid and translocation-negative subtypes of pediatric leukemia. *Cancer Res.*, **64,** 2000-2006.

25. Agirre, X., Novo, F.J., Calasanz, M.J., Larrayoz, M.J., Lahortiga, I., Valganon, M., Garcia-Delgado, M. and Vizmanos, J.L. (2003) *TP53* is frequently altered by methylation, mutation, and/or deletion in acute lymphoblastic leukaemia. *Mol. Carcinog.*, **38,** 201-208.

26. Gutierrez, M.I., Siraj, A.K., Bhargava, M., Ozbek, U., Banavali, S., Chaudhary, M.A., El Solh, H. and Bhatia, K. (2003) Concurrent methylation of multiple genes in childhood ALL: Correlation with phenotype and molecular subgroup. *Leukemia,* **17,** 1845-1850.

27. Agirre, X., Vizmanos, J.L., Calasanz, M.J., Garcia-Delgado, M., Larrayoz, M.J. and Novo, F.J. (2003) Methylation of CpG dinucleotides and/or CCWGG motifs at the promoter of TP53 correlates with decreased gene expression in a subset of acute lymphoblastic leukemia patients. *Oncogene*, **22,** 1070-1072.

28. Garcia-Manero, G., Jeha, S., Daniel, J., Williamson, J., Albitar, M., Kantarjian, H.M. and Issa, J.P. (2003) Aberrant DNA methylation in pediatric patients with acute lymphocytic leukemia. *Cancer*, **97,** 695-702.

29. Chen, H. and Wu, S. (2002) Hypermethylation of the *p15*(*INK4B*) gene in acute leukemia and myelodysplastic syndromes. *Chin. Med. J.* (Engl), **115**, 987-990.

30. Kuerbitz, S.J., Pahys, J., Wilson, A., Compitello, N. and Gray, T.A. (2002) Hypermethylation of the imprinted *NNAT* locus occurs frequently in pediatric acute leukemia. *Carcinogenesis,* **23,** 559-564.

31. Roman-Gomez, J., Castillejo, J.A., Jimenez, A., Gonzalez, M.G., Moreno, F., Rodriguez Mdel, C., Barrios, M., Maldonado, J. and Torres, A. (2002) 5' CpG island hypermethylation is associated with transcriptional silencing of the $p21^{CIP1/WAF1/SDI1}$ gene and confers poor prognosis in acute lymphoblastic leukemia. *Blood,* **99**, 2291-2296.

32. Liu, M., Taketani, T., Li, R., Takita, J., Taki, T., Yang, H.W., Kawaguchi, H., Ida, K., Matsuo, Y. and Hayashi, Y. (2001) Loss of *p73* gene expression in lymphoid leukemia cell lines is associated with hypermethylation. *Leuk. Res.*, **25,** 441-447.

33. Corn, P.G., Smith, B.D., Ruckdeschel, E.S., Douglas, D., Baylin, S.B. and Herman, J.G. (2000) E-cadherin expression is silenced by 5' CpG island methylation in acute leukemia. *Clin. Cancer Res.*, **6,** 4243-4248.

34. Guo, S.X., Taki, T., Ohnishi, H., Piao, H.Y., Tabuchi, K., Bessho, F., Hanada, R., Yanagisawa, M. and Hayashi, Y. (2000) Hypermethylation of *p16* and *p15* genes and RB protein expression in acute leukemia. *Leuk. Res.*, **24,** 39-46.

35. Cameron, E.E., Baylin, S.B. and Herman, J.G. (1999) $p15^{INK4B}$ CpG island methylation in primary acute leukemia is heterogeneous and suggests density as a critical factor for transcriptional silencing. *Blood,* **94**, 2445-2451.

36. Corn, P.G., Kuerbitz, S.J., van Noesel, M.M., Esteller, M., Compitello, N., Baylin, S.B. and Herman, J.G. (1999) Transcriptional silencing of the *p73* gene in acute lymphoblastic leukemia and Burkitt's lymphoma is associated with 5' CpG island methylation. *Cancer Res.*, **59,** 3352-3356.

37. Nakamura, M., Sugita, K., Inukai, T., Goi, K., Iijima, K., Tezuka, T., Kojika, S., Shiraishi, K., Miyamoto, N., Karakida, N. *et al.* (1999) *p16/MTS1/INK4A* gene is frequently inactivated by hypermethylation in childhood acute lymphoblastic leukemia with 11q23 translocation. *Leukemia,* **13**, 884-890.

38. Iravani, M., Dhat, R. and Price, C.M. (1997) Methylation of the multi tumor suppressor gene-2 (MTS2, CDKN1, $p15^{INK4B}$) in childhood acute lymphoblastic leukemia. *Oncogene,* **15,** 2609-2614.

39. Leegwater, P.A.J., Lambooy, L.H.J., De Abreu, R.A., Bökkerink, J.P.M. and van den Heuvel L.P. (1997) DNA methylation patterns in the calcitonin gene region at first diagnosis and at relapse of acute lymphoblastic leukemia (ALL). *Leukemia*, **11**, 971-978.

40. Herman, J.G., Jen, J., Merlo, A. and Baylin, S.B. (1996) Hypermethylation-associated inactivation indicates a tumor suppressor role for $p15^{INK4B}$. *Cancer Res.*, **56**, 722-727.

41. Hosack, D.A., Dennis, Jr G., Sherman, B.T., Lane, H.C. and Lempicki, R.A. (2003) Identifying biological themes within lists of genes with EASE. *Genome Biol.*, **4**, R70.

Table S3. Hypermethylated genes in pediatric B-lineage ALL[a]

Reference	Method	No. of genes investigated	No. of genes hypermethylated	Genes validated to be hypermethylated
Paulsson *et al.* (1)	Multiplex ligation-dependent probe amplification + bisulfite sequencing	35	16	*APC, CADM1, CD44, CDH13, CDKN2B, CHFR, ESR1, FHIT, GATA5, MSH6, PAX6, RARB, THBS1, TMP3, TP73* and *WT1*
Taylor *et al.* (2)	454 bisulfite sequencing	25	Not given	Not given
Taylor *et al.* (3)	Microarray + bisulfite sequencing	2,267	262	*ABCB1*, *DCC*, *DDX51*, *DLC1*, *KCNK2*, *LRP1B*, *NKX6-1*, *NOPE*, *PCDHGA12*, *RUNDC3B* and *SLC2A14*
Roman-Gomez *et al.* (4)	Bisulfite sequencing	1	1	*WNT5A*
Cheng *et al.* (5)	Bisulfite sequencing	1	1	*GGH*
Roman-Gomez *et al.* (6)	Bisulfite sequencing	38	38	*ADAMTS1, ADAMTS5, APAF1, CDH1, CDH13, CDKN1C, CDKN2A*, *CDKN2B, DACT1, DAPK1, DBC1, DIABLO, DKK3, FHIT, KLK10, LATS1, LATS2, P14ARF, PACRG, PARK2, PGR, PPP1R13B, PRDM2, PTEN, PTPN6, PYCARD, RPRM, SEPT4, SFRP1, SFRP2, SFRP4, SFRP5, SLC19A1, SMC1A, SMC1B, SYK, TP73* and *WIF1*
Pãixao *et al.* (7)	Bisulfite sequencing	6	6	*CALCA, CDH1, CDKN2A, GSTP1, HIC1* and *TP73*
Stam *et al.* (8)	Bisulfite sequencing	1	1	*FHIT*
Agirre *et al.* (9)	Bisulfite sequencing	2	2	*PACRG* and *PARK2*

Yang *et al.* (10)	Bisulfite sequencing	14	4	*CDKN2A*, *CDKN2B*, *FHIT* and *RARB*
Canalli *et al.* (11)	Bisulfite sequencing	3	3	*CDKN1C, CDKN2B* and *TP73*
Tsellou *et al.* (12)	Bisulfite sequencing	1	1	*CDKN2B*
Sahu and Das (13)	Bisulfite sequencing	1	1	*TP73*
Scholz *et al.* (14)	Microarray + bisulfite sequencing	57	5	*CCND2, CDC2, EGR4, MOS* and *TUSC3*
Roman-Gomez *et al.* (15)	Bisulfite sequencing	1	1	*DKK3*
Zheng *et al.* (16)	Bisulfite sequencing	1	1	*FHIT*
Gutierrez *et al.* (17)	Bisulfite sequencing	7	6	*CDH1, DAPK1, MGMT, TP73, CDKN2A* and *CDKN2B*
Garcia-Manero *et al.* (18)	Bisulfite sequencing	7	5	*ABCB1, ABL1, CDKN2B, ESR1* and *MME*
Roman-Gomez *et al.* (19)	Restriction digestion	1	1	*CDKN1A*
Corn *et al.* (20)	Bisulfite sequencing	1	1	*TP73*
Iravani *et al.* (21)	Southern blot	2	1	*CDKN2B*

[a]Of the 40 publications on hypermethylated genes in B-lineage ALL (Supplementary Material, Table S2), 21 included samples from pediatric ALL.

References

1. Paulsson, K., An, Q., Moorman, A.V., Parker, H., Molloy, G., Davies, T., Griffiths, M., Ross, F.M., Irving, J., Harrison, C.J. *et al.* (2009) Methylation of tumour suppressor gene promoters in the presence and absence of transcriptional silencing in high hyperdiploid acute lymphoblastic leukaemia. *Br. J. Haematol.*, **144**, 838-847.

2. Taylor, K.H., Kramer, R.S., Davis, J.W., Guo, J., Duff, D.J., Xu, D., Caldwell, C.W. and Shi, H. (2007) Ultradeep bisulfite sequencing analysis of DNA methylation patterns in multiple gene promoters by 454 sequencing. *Cancer Res.*, **67**, 8511-8518.

3. Taylor, K.H., Pena-Hernandez, K.E., Davis, J.W., Arthur, G.L., Duff, D.J., Shi, H., Rahmatpanah, F.B., Sjahputera, O. and Caldwell, C.W. (2007) Large-scale CpG methylation analysis identifies novel candidate genes and reveals methylation hotspots in acute lymphoblastic leukemia. *Cancer Res.*, **67**, 2617-2625.

4. Roman-Gomez, J., Jimenez-Velasco, A., Cordeu, L., Vilas-Zornoza, A., San Jose-Eneriz, E., Garate, L., Castillejo, J.A., Martin, V., Prosper, F., Heiniger, A. *et al.* (2007) *WNT5A*, a putative tumour suppressor of lymphoid malignancies, is inactivated by aberrant methylation in acute lymphoblastic leukaemia. *Eur. J. Cancer*, **43**, 2736-2746.

5. Cheng, Q., Cheng, C., Crews, K.R., Ribeiro, R.C., Pui, C.H., Relling, M.V. and Evans, W.E. (2006) Epigenetic regulation of human γ-glutamyl hydrolase activity in acute lymphoblastic leukemia cells. *Am. J. Hum. Genet.*, **79**, 264-274.

6. Roman-Gomez, J., Jimenez-Velasco, A., Agirre, X., Castillejo, J.A., Navarro, G., Calasanz, M.J., Garate, L., San Jose-Eneriz, E., Cordeu, L., Prosper, F. *et al.* (2006) CpG island methylator phenotype redefines the prognostic effect of t(12;21) in childhood acute lymphoblastic leukemia. *Clin. Cancer Res.*, **12**, 4845-4850.

7. Pãixao, V.A., Vidal, D.O., Caballero, O.L., Vettore, A.L., Tone, L.G., Ribeiro, K.B. and Lopes, L.F. (2006) Hypermethylation of CpG island in the promoter region of *CALCA* in acute lymphoblastic leukemia with central nervous system (CNS) infiltration correlates with poorer prognosis. *Leuk. Res.*, **30**, 891-894.

8. Stam, R.W., den Boer, M.L., Passier, M.M., Janka-Schaub, G.E., Sallan, S.E., Armstrong, S.A. and Pieters, R. (2006) Silencing of the tumor suppressor gene *FHIT* is highly characteristic for *MLL* gene rearranged infant acute lymphoblastic leukemia. *Leukemia,* **20,** 264-271.

9. Agirre, X., Román-Gómez, J., Vázquez, I., Jiménez-Velasco, A., Garate, L., Montiel-Duarte, C., Artieda, P., Cordeu, L., Lahortiga, I., Calasanz, M.J. *et al.* (2006) Abnormal methylation of the common *PARK2* and *PACRG* promoter is associated with downregulation of gene expression in acute lymphoblastic leukemia and chronic myeloid leukemia. *Int. J. Cancer*, **118**, 1945-1953.

10. Yang, Y., Takeuchi, S., Hofmann, W.K., Ikezoe, T., van Dongen, J.J., Szczepanski, T., Bartram, C.R., Yoshino, N., Taguchi, H. and Koeffler, H.P. (2006) Aberrant methylation in promoter-associated CpG islands of multiple genes in acute lymphoblastic leukemia. *Leuk. Res.*, **30,** 98-102.

11. Canalli, A.A., Yang, H., Jeha, S., Hoshino, K., Sanchez-Gonzalez, B., Brandt, M., Pierce, S., Kantarjian, H., Issa, J.P. and Garcia-Manero, G. (2005) Aberrant DNA methylation of a cell cycle regulatory pathway composed of P73, P15 and P57KIP2 is a rare event in children with acute lymphocytic leukemia. *Leuk. Res.*, **29,** 881-885.

12. Tsellou, E., Troungos, C., Moschovi, M., Athanasiadou-Piperopoulou, F., Polychronopoulou, S., Kosmidis, H., Kalmanti, M., Hatzakis, A., Dessypris, N., Kalofoutis, A. *et al.* (2005) Hypermethylation of CpG islands in the promoter region of the *p15INK4B* gene in childhood acute leukaemia. *Eur. J. Cancer*, **41,** 584-589.

13. Sahu, G.R. and Das, B.R. (2005) Alteration of p73 in pediatric de novo acute lymphoblastic leukemia. *Biochem. Biophys. Res. Commun.*, **327**, 750-755.

14. Scholz, C., Nimmrich, I., Burger, M., Becker, E., Dorken, B., Ludwig, W.D. and Maier, S. (2005) Distinction of acute lymphoblastic leukemia from acute myeloid leukemia through microarray-based DNA methylation analysis. *Ann. Hematol.*, **84,** 236-244.

15. Roman-Gomez, J., Jimenez-Velasco, A., Agirre, X., Castillejo, J.A., Navarro, G., Barrios, M., Andreu, E.J., Prosper, F., Heiniger, A. and Torres, A. (2004) Transcriptional silencing of the *Dickkopfs-3* (*Dkk-3*) gene by CpG hypermethylation in acute lymphoblastic leukaemia. *Br. J. Cancer,* **91,** 707-713.

16. Zheng, S., Ma, X., Zhang, L., Gunn, L., Smith, M.T., Wiemels, J.L., Leung, K., Buffler, P.A. and Wiencke, J.K. (2004) Hypermethylation of the 5' CpG island of the *FHIT* gene is associated with hyperdiploid and translocation-negative subtypes of pediatric leukemia. *Cancer Res.*, **64,** 2000-2006.

17. Gutierrez, M.I., Siraj, A.K., Bhargava, M., Ozbek, U., Banavali, S., Chaudhary, M.A., El Solh, H. and Bhatia, K. (2003) Concurrent methylation of multiple genes in childhood ALL: Correlation with phenotype and molecular subgroup. *Leukemia,* **17,** 1845-1850.

18. Garcia-Manero, G., Jeha, S., Daniel, J., Williamson, J., Albitar, M., Kantarjian, H.M. and Issa, J.P. (2003) Aberrant DNA methylation in pediatric patients with acute lymphocytic leukemia. *Cancer*, **97,** 695-702.

19. Roman-Gomez, J., Castillejo, J.A., Jimenez, A., Gonzalez, M.G., Moreno, F., Rodriguez Mdel, C., Barrios, M., Maldonado, J. and Torres, A. (2002) 5' CpG island hypermethylation is associated with transcriptional silencing of the $p21^{CIP1/WAF1/SDI1}$ gene and confers poor prognosis in acute lymphoblastic leukemia. *Blood,* **99**, 2291-2296.

20. Corn, P.G., Kuerbitz, S.J., van Noesel, M.M., Esteller, M., Compitello, N., Baylin, S.B. and Herman, J.G. (1999) Transcriptional silencing of the *p73* gene in acute lymphoblastic leukemia and Burkitt's lymphoma is associated with 5' CpG island methylation. *Cancer Res*., **59,** 3352-3356.

21. Iravani, M., Dhat, R. and Price, C.M. (1997) Methylation of the multi tumor suppressor gene-2 (MTS2, CDKN1, $p15^{INK4B}$) in childhood acute lymphoblastic leukemia. *Oncogene,* **15,** 2609-2614.

Table S4. Chromosomal distribution of the 8,662 genes with significant CpG island methylation scores

Chromosome	No. of genes only hypermethylated in *ETV6/RUNX1*-positive cases	No. of genes only hypermethylated in high hyperdiploid cases	No of. genes hypermethylated in both genetic subgroups	Total no. of hypermethylated genes	Total no. of genes hypermethylated in *ETV6/RUNX1*-positive cases	Total no. of genes hypermethylated in high hyperdiploid cases
1	88	453	261	802	349	714
2	83	281	189	553	272	470
3	49	224	128	401	177	352
4	43	211	121	375	164	332
5	62	184	119	365	181	303
6	46	300	135	481	181	435
7	46	191	180	417	226	371
8	0	282	0	282	0	282
9	60	173	131	364	191	304
10	25	178	121	324	146	299
11	62	327	174	563	236	501
12	46	249	116	411	162	365
13	18	64	58	140	76	122
14	24	157	76	257	100	233
15	39	122	85	246	124	207
16	61	160	145	366	206	305
17	60	229	176	465	237	404

18	16	77	52	145	68	129
19	83	255	236	574	320	490
20	30	120	88	238	118	208
21	8	80	61	149	69	141
22	23	88	86	197	110	173
X	31	272	198	501	227	472
Y	9	2	35	46	43	38
Total	1012	4679	2971	8662	3983	7650

Table S5. The 30 genes displaying the highest methylation peak scores among the 8,662 genes with significant CpG island methylation scores

Gene[a]	Function[b]	Location[b]
DNAJB8	Heat shock protein	3q21.3
LSP1	Cell motility	11p15.5
DPCR1	Unknown	6p21.33
GLI2	Transcription regulation	2q14
ZFHX4	Transcription regulation	8q21.11
SLC5A3	Ion transport	21q22.12
SLC44A5	Unknown	1p31.1
AQP6	Transport	12q13
SSX7	Transcription regulation	Xp11.23
ATP9B	Metabolism	18q23
CD7	Immune response	17q25.2-25.3
OTUD4	Unknown	4q31.21
FTSJ1	Methyltranferase	Xp11.23
GJD4	Cell signaling	10p11.21
ANGPT2	Angiogenesis	8p23.1
HIST1H2AA	Histone	6p22.2
HIST1H2BA	Histone	6p22.2
WISP3	Cell signaling	6q21
MOBP	Unknown	3p22.1
LRRC37A	Unknown	17q21.31
AFAP1	Morphogenesis	4p16
ATP11A	Metabolism	13q34
SSX3	Transcription regulation	Xp11.23
GAS7	Cell cycle	17p13.1
RBMY1J	RNA processing	Yq11.22
IQSEC3	Signal transduction	12p13.33
SSX2	Unknown	Xp11.22
TULP4	Transcription regulation	6q25-26
MAGEB10	Unknown	Xp21.3
PHOX2B	Transcription regulation	4p12

[a]The genes are listed in order of decreasing -log10 p-values (see Materials and Methods).

[b]Function and location are based on data in NCBI Entrez Gene (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene). Only one major function per gene is given.

Table S6. The 30 genes displaying the highest methylation peak scores in *ETV6/RUNX1*-positive and high hyperdiploid cases

ETV6/RUNX1-positive cases			High hyperdiploid cases		
Gene[a]	Function[b]	Location[b]	Gene[a]	Function[b]	Location[b]
OTUD4	Unknown	4q31.21	*REXO1L1*	Exonuclease	8q21.2
ATP9B	Metabolism	18q23	*SLC17A4*	Ion transport	6p21.3-22
GLI2	Transcription regulation	2q14	*DEFB109*	Unknown	8p23.1
MOBP	Unknown	3p22.1	*DNAJB8*	Heat shock protein	3q21.3
LRRC37A	Unknown	17q21.31	*SPRYD5*	Unknown	11q11
RBMY1J	RNA processing	Yq11.22	*UGT2B11*	Metabolism	4q13.2
SSX2	Transcription regulation	Xp11.22	*PZP*	Pregnancy	12p12.2-13
TULP4	Transcription regulation	6q25-26	*DPCR1*	Unknown	6p21.33
LSP1	Cell motility	11p15.5	*SYT16*	Unknown	14q23.2
LTBP1	Unknown	2p21-22	*GLI2*	Transcription regulation	2q14
MAGEB10	Unknown	Xp21.3	*MAPK10*	Cell signaling	4q22.1-23
SSX3	Transcription regulation	Xp11.23	*LEMD1*	Unknown	1q32.1
PHOX2B	Transcription regulation	4p12	*ZFHX4*	Transcription regulation	8q21.11
CD7	Immune response	17q25.2-25.3	*SLC5A3*	Ion transport	21q22.12
ATP11A	Transport	13q34	*SLC44A5*	Unknown	1p31.1
RAB27A	Cell signaling	15q15-21.1	*AQP6*	Transport	12q13
CPSF3L	snRNA regulation	1p36.33	*CD7*	Immune response	17q25.2-25.3
TAS1R3	Cell signaling	1p36.33	*ATP9B*	Metabolism	18q23
KCND2	Transport	7q31	*FTSJ1*	Methyltranferase	Xp11.23

TRAF7	Transcription regulation	16p13.3	*GJD4*	Cell signaling	10p11.21
HOXA3	Transcription regulation	7p14-15	*ANGPT2*	Angiogenesis	8p23.1
HOXA4	Transcription regulation	7p14-15	*HIST1H2AA*	Histone	6p22.2
PAGE2B	Unknown	Xp11.21	*HIST1H2BA*	Histone	6p22.2
ZNF609	Unknown	15q22.31	*TRDN*	Muscle function	6q22-23
UBE2U	Protein degradation	1p31.3	*WISP3*	Cell signaling	6q21
CSF2RA	Cell signaling	Xp22.32, Yp11.3	*AFAP1*	Morphogenesis	4p16
TSPY1	Cell proliferation	Yp11.2	*ATP11A*	Transport	13q34
TSPY2	Cell differentiation	Yp11.2	*AMELX*	Mineralization	Xp22.1-22.31
DNAJB8	Heat shock protein	3q21.3	*CLCN5*	Transport	Xp11.23
HIST1H2AA	Histone	6p22.2	*AHRR*	Cell signaling	5p15.3

[a]The genes are listed in order of decreasing -log10 p-values (see Methods section).

[b]Function and location are based on data in NCBI Entrez Gene (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene). Only one major function per gene is given.

Table S7. The 30 unique genes displaying the highest methylation peak scores in *ETV6/RUNX1*-positive and high hyperdiploid cases

ETV6/RUNX1-positive cases			High hyperdiploid cases		
Gene[a]	Function[b]	Location[b]	Gene[a]	Function[b]	Location[b]
OTUD4	Unknown	4q31.21	*KLHDC7B*	Unknown	22q13.33
SSX7	Transcription regulation	Xp11.23	*GP1BB*	Signal transduction	22q11.21
BDNF	Growth factor	11p13	*BRP44L*	Unknown	6q27
LSP1	Cell motility	11p15.5	*GZF1*	Transcription regulation	20p11.21-12.3
MS4A12	Cell signaling	11q12	*MFSD3*	Unknown	8q24.3
PTPN3	Cell signaling	9q31	*TMEM200B*	Unknown	1p35
SSX3	Transcription regulation	Xp11.23	*CSAG1*	Tumor antigen	Xq28
GGTL4	Catalytic activity	22q11.22	*MAGEA12*	Unknown	Xq28
SSX2	Transcription regulation	Xp11.22	*TCEB3C*	Transcription regulation	18q21.1
SYNE2	Morphogenesis	14q23.2	*RPL36*	Translation	19p13.3
CPN1	Proteolysis	10q24.2	*FGF20*	Cell growth	8p21.3-22
LRRC37A	Unknown	17q21.31	*GAS2*	Apoptosis	11p14.3-15.2
SPANXB2	Spermatoid development	Xq27.1	*GNL2*	GTPase	1p34.3
ANP32D	Tumor suppressor	12q13.11	*LCK*	Hematopoiesis	1p34.3
PCDHA12	Cell adhesion	5q31	*PGA5*	Unknown	11q13
TSPY1	Cell proliferation	Yp11.2	*SLC25A25*	Unknown	9q34.11
TSPY2	Cell proliferation	Yp11.2	*TNS3*	Cell proliferation	7p12.3
NALP4	Apoptosis	19q13.42	*TRPV1*	Ion transport	17p13.3
FAM156A	Unknown	Xp11.22	*AQP5*	Water transport	12q13

TP53TG3	Unknown	16p13	*C4A*	Inflammatory response	6p21.3
RBMY1A1	RNA binding	Yq11.22	*C4B*	Inflammatory response	6p21.3
RBMY1B	RNA binding	Yq11.22	*NCAN*	Cell adhesion	19p12
RBMY1F	RNA binding	Yq11.22	*EIF4G1*	Translation	3q27-29
ANAPC11	Cell cycle	17q25.3	*FHL2*	Transcription regulation	2q12-14
GPR109B	Signal transduction	12q24.31	*TMEM176A*	Unknown	7q36.1
FOXD4L1	Transcription regulation	2q14.1	*HEXDC*	Unknown	17q25.3
HOXA3	Transcription regulation	7p14-15	*TMEM176B*	Morphogenesis	7q36.1
IQSEC3	Signal transduction	12p13.33	*LYPD5*	Unknown	19q13.31
SPANXE	Spermatid development	Xq27.1	*ODF3*	Sperm component	11p15.5
HOXA4	Transcription regulation	7p14-15	*RPUSD1*	Unknown	16p13.3

[a]The genes are listed in order of decreasing -log10 p-values (see Methods section).

[b]Function and location are based on data in NCBI Entrez Gene (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene). Only one major function per gene is given.

Table S8. The genes shown by segmental analysis to be highly enriched for methylation but that did not display decreased expression or for which no expression data were available. The 45 genes associated with decreased expression are given in Table 2 in the article.

Gene	Function[a]	Location[a]	Expression
FOXA2	Transcription regulation	20p11	Not available
FOXD4L3	Transcription regulation	9q13	Not available
FUSSEL18	Transcription regulation	18q21.1	Not available
GLIS1	Transcription regulation	1p32.3	Not available
HDAC9	Transcription regulation	7p21.1	Not available
PHOX2B	Transcription regulation	4p12	Not available
SNAPC4	Transcription regulation	9q34.3	Not available
SSBP3	Transcription regulation	1p32.3	Not available
SSX2	Transcription regulation	Xp11.22	Not available
SSX5	Transcription regulation	Xp11.23	Not available
SSX7	Transcription regulation	Xp11.23	Not available
YY2	Transcription regulation	Xp22.1-22.2	Not available
ZNF384	Transcription regulation	12p12	Not available
ZNF502	Transcription regulation	3p21.31	Not available
CAMK1G	Cell signaling	1q32-41	Not available
CCR5	Cell signaling	3p21.31	Not available
ERAS	Cell signaling	Xp11.23	Not available
FYN	Cell signaling	6q21	Not available
GDF9	Cell signaling	5q31.1	Not available
ICOS	Cell signaling	2q33	Not available
IL2	Cell signaling	4q26-27	Not available
MS4A8B	Cell signaling	11q12.2	Not available
TLR3	Cell signaling	4q35	Not available
TRPV1	Cell signaling	17p13.3	Not available
WISP3	Cell signaling	6q21	Not available
AFAP1	Apoptosis	4p16	Not available
CASP5	Apoptosis	11q22.2-22.3	Not available
HUS1B	Apoptosis	6p25.3	Not available
ICEBERG	Apoptosis	11q21-22	Not available
MAGED1	Apoptosis	Xp11.23	Not available
SYNE1	Apoptosis	6q25	Not available

TRAF7	Apoptosis	16p13.3	Not available
CTNNA3	Cell adhesion	10q22.2	Not available
DSC2	Cell adhesion	18q12.1	Not available
LSP1	Cell adhesion	11p15.5	Not available
HIST1H1B	Nucleosome assembly	6p21.3-22	Not available
HIST1H2AA	Nucleosome assembly	6p22.2	Not available
HIST1H3E	Nucleosome assembly	6p21.3	Not available
CABLES1	Cell cycle	18q11.2	Not available
DMBT1	Cell cycle	10q25.3-26.1	Not available
CD7	Immune response	17q25.2-25.3	Not available
FCRL3	Immune response	1q21-22	Not available
DCN	Morphogenesis	12q21.33	Not available
ITIH5	Morphogenesis	10p14	Not available
MAS1	Cell proliferation	6q25.3-26	Not available
TNS3	Cell proliferation	7p12.3	Not available
ELA2B	Proteolysis	1p36.21	Not available
DNPEP	Proteolysis	2q35	Not available
B3GALT2	Glycosylation	1q31	Not available
CES1	Metabolism	16q13-22.1	Not available
CMTM5	Chemotaxis	14q11.2	Not available
IL21	Cell maturation	4q26-27	Not available
SETD2	Chromatin modification	3p21.31	Not available
SLC44A5	Transport	1p31.1	Not available
TRPC7	Ion transport	5q31.2	Not available
DNTT	Various	10q23-24	Not available
GLI2	Various	2q14	Not available
MATK	Various	19p13.3	Not available
RAG2	Various	11p13	Not available
ANP32D	Unknown	12q13.11	Not available
DIRC1	Unknown	2q33	Not available
LEMD1	Unknown	1q32.1	Not available
REG4	Unknown	1p12-13.1	Not available
CD69	Cell signaling	12p12-13	Not differential
GJB1	Cell signaling	Xq13.1	Not differential
LYNX1	Cell signaling	8q24.3	Not differential
PTPRF	Cell signaling	1p34	Not differential

PTPRU	Cell signaling	1p35.1-35.3	Not differential
RGS10	Cell signaling	10q25	Not differential
IL1A	Apoptosis	2q14	Not differential
PRMT2	Apoptosis	21q22.3	Not differential
EWSR1	Transcription regulation	22q12.2	Not differential
SSX3	Transcription regulation	Xp11.23	Not differential
APTX	DNA repair	9p13.3	Not differential
ATP2B2	Ion transport	3p25.3	Not differential
ATRX	Methylation regulation	Xq13.1-21.1	Not differential
CD53	Signal transduction	1p13	Not differential
DDX59	ATP binding	1q32.1	Not differential
DNAJB8	Protein folding	3q21.3	Not differential
RBMY1A1	Cell differentiation	Yq11.22	Not differential
SELE	Cell adhesion	1q22-25	Not differential
TCL6	Unknown	14q32.1	Not differential
MLL3	Transcription regulation	7q36.1	Upregulated
TIGD2	Transcription regulation	4q22.1	Upregulated
FGFR1	Cell signaling	8p11.1-11.2	Upregulated
IL2RG	Cell signaling	Xq13.1	Upregulated
ANGPTL7	Signal transduction	1p36.2-36.3	Upregulated
KIFAP3	Signal transduction	1q24.2	Upregulated
ALKBH1	DNA repair	14q24.3	Upregulated
CTNNA1	Transport	5q31	Upregulated
LRMP	Vesicle function	12p12.1	Upregulated
MS4A1	B-cell differentiation	11q12	Upregulated
NEK7	Mitotic control	1q31.3	Upregulated

[a]Function and location are based on data in NCBI Entrez Gene (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene). Only one major function per gene is given.

Table S9. Genes involved in B-cell development and regulation[a]

Gene	Function[b]	Location[c]
ABL1	Cell signaling	9q34.1
BAD	Apoptosis	11q13.1
BCL2	Apoptosis	18q21.3
BCL2L11	Apoptosis	2q13
BCL6	Transcription regulation	3q27
BCL10	Apoptosis	1p22
BCL11A	Transcription regulation	2p16.1
BLK	Cell signaling	8p22-23
BLNK	Cell signaling	10q23.2-23.3
BTK	Apoptosis	Xq21.3-22
BTLA	Immune response	3q13.2
CD19	Cell signaling	16p11.2
CD79A	Cell signaling	19q13.2
CD79B	Cell signaling	17q23
CD40LG	Immune response	Xq26
CSF1R	Cell signaling	5q33-35
DNTT	DNA synthesis	10q23-24
EBF1	Transcription regulation	5q34
FCGR2B	Immune response	1q23
FLT3	Cell signaling	13q12
FLT3LG	Cell signaling	19q13.3
IFNAR1	Cell signaling	21q22.1
IFNG	Immune response	12q14
IGH@	Immune response	14q32.3
IGK@	Immune response	2p12
IGL@	Immune response	22q11.1-11.2
IKBKB	Transcription regulation	8p11.2
IKZF1	Transcription regulation	7p11.1-13
IKZF2	Transcription regulation	2q34
IKZF3	Transcription regulation	17q21
IL2RA	Immune response	10p14-15
IL2RB	Immune response	22q13.1
IL2RG	Immune response	Xq13.1

IL4R	Immune response	16p11.2-12.1
IL7R	Immune response	5p13
IRF4	Transcription regulation	6p23-25
IRF8	Transcription regulation	16q24.1
LAG3	Immune response	12p13.3
LAX1	Immune response	1q32.1
LEF1	Cell signaling	4q23-25
LILRB4	Immune response	19q13.4
LYN	Cell signaling	8q13
NFKB1	Transcription regulation	4q24
NFKB2	Transcription regulation	10q24
NOTCH1	Cell signaling	9q34.3
NOTCH2	Cell cycle	1p11-13
PAX5	Transcription regulation	9p13
PLCG2	Cell signaling	16q24.1
RAG1	DNA recombination	11p13
RAG2	DNA recombination	11p13
SOX4	Transcription regulation	6p22.3
SPI1	Transcription regulation	11p11.2
SPIB	Transcription regulation	19q13.3-13.4
SYK	Cell signaling	9q22
TCF3	Transcription regulation	19p13.3
VPREB1	Immune response	22q11.22

[a]In order to identify putative targets for hypermethylation in B-lineage ALL, we compiled genes known to be involved in B-cell development and regulation using the pathway analysis software provided by Ingenuity Systems (Ingenuity Systems, Redwood, CA) and, to a lesser extent, PubMed. A total of 56 genes strongly associated with these features were identified.

[b]Function is based on data produced by the EASE software (1) or found in NCBI Entrez Gene (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene). Only one major function per gene is given.

[c]Location is based on information in NCBI Entrez Gene (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene).

Reference

1. Hosack, D.A., Dennis, Jr G., Sherman, B.T., Lane, H.C. and Lempicki, R.A. (2003) Identifying biological themes within lists of genes with EASE. *Genome Biol.*, **4**, R70.

Table S10. Genes involved in gene fusions in B-lineage ALL[a]

Gene	Function[b]	Location[c]
ABL1	Cell signaling	9q34.1
AFF1	Transcription regulation	4q21
AFF3	Transcription regulation	2q11.2-12
AFF4	Transcription regulation	5q31
BAZ2A	Transcription regulation	12q24.3
BCL2	Apoptosis	18q21.3
BCL6	Transcription regulation	3q27
BCL7A	Transcription regulation	12q24.13
BCL9	Cell signaling	1q21
BCR	Cell signaling	22q11.2
CDKN2A	Cell cycle	9p21
CEBPA	Transcription regulation	19q13.1
CEBPB	Transcription regulation	20q13.1
CEBPD	Transcription regulation	8q11.2
CEBPE	Transcription regulation	14q11.2
CEBPG	Transcription regulation	19q13.11
CREBBP	Transcription regulation	16p13.3
DAZAP1	Morphogenesis	19p13.3
DDX6	RNA binding	11q23.3
ELN	Morphogenesis	7q11.2
ERG	Transcription regulation	21q22.3
ETV6	Transcription regulation	12p13
EWSR1	Transcription regulation	22q12.2
FGFR1	Cell signaling	8p11.1-p11.2
FGFR1OP	Cell proliferation	6q27
FOXP1	Transcription regulation	3p14.1
FRYL	Transcription regulation	4p12
FUS	RNA binding	16p11.2
GAS7	Cell cycle	17p13.1
HLF	Transcription regulation	17q22
ID4	Transcription regulation	6p21-22
IGH@	Immune response	14q32.3
IGK@	Immune response	2p12
IGL@	Immune response	22q11.1-11.2

IL3	Cell signaling	5q31.1
JAK2	Cell signaling	9p24
LHX4	Transcription regulation	1q25.2
MAPRE1	Cell cycle	20q11.1-11.2
MEF2D	Transcription regulation	1q12-23
MLL	Transcription regulation	11q23
MLLT1	Transcription regulation	19p13.3
MLLT4	Cell signaling	6q27
MLLT10	Transcription regulation	10p12
MYC	Transcription regulation	8q24.2
PAX5	Transcription regulation	9p13
PBX1	Transcription regulation	1q23
PCM1	Morphogenesis	8p21.3-22
PML	Transcription regulation	15q22
POU2AF1	Transcription regulation	11q23.1
PVT1	Unknown	8q24
RANBP17	Transport	5q34
RUNX1	Transcription regulation	21q22.3
SFPQ	RNA processing	1p34.3
STL	Unknown	6q22
TAF15	Transcription regulation	17q11.1-11.2
TCF3	Transcription regulation	19p13.3
TFPT	Transcription regulation	19q13
TTL	Unknown	13q14.11
ZNF384	Transcription regulation	12p12
ZNF521	Unknown	18q11.2

[a]Genes known to be involved in gene fusions in B-lineage ALL may also be putative targets for differential methylation, and all 60 such genes were retrieved from the Mitelman Database of Chromosome Aberrations in Cancer (http://cgap.nci.nih.gov/Chromosomes/Mitelman).

[b]Function is based on data produced by the EASE software (1) or found in NCBI Entrez Gene (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene). Only one major function per gene is given.

[c]Location is based on information in NCBI Entrez Gene (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene).

Reference

1. Hosack, D.A., Dennis, Jr G., Sherman, B.T., Lane, H.C. and Lempicki, R.A. (2003) Identifying biological themes within lists of genes with EASE. *Genome Biol*., **4**, R70.

Table S11. Genes mutated in B-lineage ALL[a]

Gene	Type of cancer gene	Location[b]
ATM	Tumor suppressor	11q22-23
BRAF	Oncogene	7q34
CDKN2A	Tumor suppressor	9p21
FLT3	Oncogene	13q12
JAK2	Oncogene	9p24
KRAS	Oncogene	12p12.1
NRAS	Oncogene	1p13.2
PTEN	Tumor suppressor	10q23.3
PTPN11	Oncogene	12q24
RUNX1	Oncogene	21q22.3
TP53	Tumor suppressor	17p13.1

[a]The Catalogue of Somatic Mutations in Cancer (COSMIC; www.sanger.ac.uk/genetics/CGP/cosmic/) was used to identify genes that have been shown to be mutated in B-lineage ALL. A total of 11 genes were found.

[b]Location is based on information in NCBI Entrez Gene (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene).

Table S12. The 131 hypermethylated genes (out of 167 shown to be recurrently methylated and/or involved in B-cell development or mutated or rearranged in B-lineage ALL) that did not display decreased expression or for which no expression data were available. The 36 genes associated with decreased expression are given in Table 3 in the article.

Gene	Function[a]	Location[a]	Expression
BCL11A	Transcription regulation	2p16.1	Not available
EGR4	Transcription regulation	2p13	Not available
HIC1	Transcription regulation	17p13.3	Not available
LHX1	Transcription regulation	17q12	Not available
MATK	Transcription regulation	19p13.3	Not available
PAX6	Transcription regulation	11p13	Not available
TAL1	Transcription regulation	1p32	Not available
TFAP2C	Transcription regulation	20q13.2	Not available
TRAF7	Transcription regulation	16p13.3	Not available
BCL7C	Apoptosis	16p11	Not available
CDC2L2	Apoptosis	1p36.3	Not available
DIABLO	Apoptosis	12q24.31	Not available
HTRA2	Apoptosis	2p12	Not available
P53AIP1	Apoptosis	11q24	Not available
SFN	Apoptosis	1p35.3	Not available
TRADD	Apoptosis	16q22	Not available
TSC22D3	Apoptosis	Xq22.3	Not available
CAV1	Transport	7q31.1	Not available
KCNK2	Transport	1q41	Not available
RANBP17	Transport	5q34	Not available
SLC5A3	Transport	21q22.12	Not available
SLC5A8	Transport	12q23.3	Not available
SLC19A1	Transport	21q22.3	Not available
SLC22A18	Transport	11p15.5	Not available
TSSC4	Transport	11p15.5	Not available
BLK	Cell signaling	8p22-23	Not available
EPHA3	Cell signaling	3p11.2	Not available
IL3	Cell signaling	5q31.1	Not available
LRIG3	Cell signaling	12q14.1	Not available
LTB4R	Cell signaling	14q11.2-12	Not available

PGR	Cell signaling	11q22-23	Not available
THEM4	Cell signaling	1q21	Not available
ATM	Cell cycle	11q22-23	Not available
CDK2AP1	Cell cycle	12q24.31	Not available
CDKN1C	Cell cycle	11p15.5	Not available
CIRBP	Cell cycle	19p13.3	Not available
KLK10	Cell cycle	19q13.3-13.4	Not available
MOS	Cell cycle	8q11	Not available
PRKCDBP	Cell cycle	11p15.4	Not available
ADAMTS1	Proteolysis	21q21.2	Not available
CST3	Proteolysis	20p11.21	Not available
DUB3	Proteolysis	8p23.1	Not available
FBXW7	Proteolysis	4q31.3	Not available
LRP1B	Proteolysis	2q21.2	Not available
MMP21	Proteolysis	10q26.2	Not available
SERPINB13	Proteolysis	18q21.3-22	Not available
GP1BB	Cell adhesion	22q11.21	Not available
PARVB	Cell adhesion	22q13.2-13.33	Not available
PCDHGA12	Cell adhesion	5q31	Not available
ADPRH	Hydrolase	3q13.31-13.33	Not available
B4GALT2	Transferase	1p33-34	Not available
DNTT	DNA replication	10q23-24	Not available
FRK	Cell proliferation	6q21-22.3	Not available
HOXD3	Morphogenesis	2q31.1	Not available
TMOD1	Morphogenesis	9q22.3	Not available
RAG2	Various	11p13	Not available
TAPBP	Various	6p21.3	Not available
H19	Unknown	11p15.5	Not available
LYPD1	Unknown	2q21.2	Not available
RHBDD3	Unknown	22q12.1-12.2	Not available
ARID3A	Transcription regulation	19p13.3	Not differential
EWSR1	Transcription regulation	22q12.2	Not differential
HOPX	Transcription regulation	4q11-12	Not differential
HOXA9	Transcription regulation	7p14-15	Not differential
ID4	Transcription regulation	6p21-22	Not differential
IRF7	Transcription regulation	11p15.5	Not differential

MBD1	Transcription regulation	18q21	Not differential
MYT1	Transcription regulation	20q13.33	Not differential
PRDM2	Transcription regulation	1p36	Not differential
RARB	Transcription regulation	3p24	Not differential
SIN3A	Transcription regulation	15q23	Not differential
SMAD4	Transcription regulation	18q21.1	Not differential
SSBP2	Transcription regulation	5q14.1	Not differential
TCF3	Transcription regulation	19p13.3	Not differential
TGIF1	Transcription regulation	18p11.3	Not differential
BAD	Apoptosis	11q13.1	Not differential
BCL2	Apoptosis	18q21.3	Not differential
CASP3	Apoptosis	4q34	Not differential
DAPK3	Apoptosis	19p13.3	Not differential
DCC	Apoptosis	18q21.3	Not differential
FASTK	Apoptosis	7q35	Not differential
GAS2	Apoptosis	11p14.3-p15.2	Not differential
LCK	Apoptosis	1p34.3	Not differential
PHF17	Apoptosis	4q26-27	Not differential
VHL	Apoptosis	3p26-25	Not differential
CADM4	Cell cycle	19q13.31	Not differential
CDKN2A	Cell cycle	9p21	Not differential
CGREF1	Cell cycle	2p23.3	Not differential
EXT1	Cell cycle	8q24.11-24.13	Not differential
FHIT	Cell cycle	3p14.2	Not differential
FOXN3	Cell cycle	14q31.3	Not differential
LIG4	Cell cycle	13q33-34	Not differential
CD19	Cell signaling	16p11.2	Not differential
DKK3	Cell signaling	11p15.2	Not differential
EDNRB	Cell signaling	13q22	Not differential
MLLT4	Cell signaling	6q27	Not differential
AXIN2	Signal transduction	17q23-24	Not differential
IGFBP1	Signal transduction	7p12-13	Not differential
LILRB4	Signal transduction	19q13.4	Not differential
PRMT2	Signal transduction	21q22.3	Not differential
CDH1	Cell adhesion	16q22.1	Not differential
SRPX	Cell adhesion	Xp21.1	Not differential

CYR61	Morphogenesis	1p22-31	Not differential
GPC3	Morphogenesis	Xq26.1	Not differential
SELENBP1	Transport	1q21-22	Not differential
SLC12A4	Transport	16q22.1	Not differential
ADAMTS5	Proteolysis	21q21.3	Not differential
CTF8	DNA replication	16q22.1	Not differential
MGAT4A	Carbohydrate processing	2q12	Not differential
NPM1	Cell proliferation	5q35	Not differential
CRMP1	Various	4p15-16.1	Not differential
IGSF8	Various	1q23.1	Not differential
IL7	Various	8q12-13	Not differential
LYNX1	Various	8q24.3	Not differential
PRKCD	Various	3p21.31	Not differential
SST	Various	3q28	Not differential
MEG3	Unknown	14q32	Not differential
MSMB	Unknown	10q11.2	Not differential
JDP2	Transcription regulation	14q24.3	Upregulated
SOX4	Transcription regulation	6p22.3	Upregulated
WT1	Transcription regulation	11p13	Upregulated
BCL2L11	Morphogenesis	2q13	Upregulated
NDRG2	Morphogenesis	14q11.2	Upregulated
NNAT	Morphogenesis	20q11.2-12	Upregulated
BLNK	Cell signaling	10q23.2-23.33	Upregulated
FGFR1	Cell signaling	8p11.1-11.2	Upregulated
CTNNA1	Apoptosis	5q31	Upregulated
GAS7	Cell cycle	17p13.1	Upregulated
TUSC3	Glycosylation	8p22	Upregulated
ADAMTS9	Various	3p14.2-14.3	Upregulated
IL2RB	Various	22q13.1	Upregulated

[a]Function and location are based on data in NCBI Entrez Gene (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene). Only one major function per gene is given.

Table S13. The genomic positions of the methylation peaks, the peak sizes, the probe densities and the positions of the peaks in relation to the transcription start sites of all genes listed Tables 2 and 3 in the manuscript

Gene	Median no. of probes per peak	Median peak size (bp)	Position of the peak center (bp)	Position of the peak center in relation to transcription start site
AR	3	239	ChrX: 66705329	Covering[a]
ARHGAP25	7	660	Chr2: 68814972	499 bp upstream
ARTN	7	559	Chr1: 44170190	579 bp upstream
BBC3	5	539	Chr19: 52426792	501 bp upstream
BCL10	3	250	Chr1: 85514796	Covering
BOK	5	459	Chr2: 242146277	571 bp upstream
BRCA1	6	529	Chr17: 38531918	924 bp upstream
BRDT	7	664	Chr1: 92187236	483 bp upstream
CASP2	5	1288	Chr7: 142695721	Covering
CASP9	5	449	Chr1: 15723633	351 bp upstream
CCND2	6	650	Chr12: 4252950	248 bp upstream
CD151	3	349	Chr11: 822388	563 bp upstream
CD2	8	539	Chr1: 117098036	588 bp upstream
CD36	8	766	Chr7: 80069158	300 bp upstream
CD44	7	444	Chr11: 35116512	480 bp upstream
CDC14B	4	454	Chr9: 98186329	3727 bp upstream
CDH13	5	444	Chr16: 81218103	Covering
CDK2AP1	6	533	Chr12: 122323116	476 bp upstream
CEBPA	3	264	Chr19 38485123	Covering
CES2	4	339	Chr16: 65525257	590 bp upstream
CLDN1	5	459	Chr3: 191523556	400 bp upstream
CPSF3L	3	359	Chr1: 1234785	1966 bp upstream
CSF1R	8	832	Chr5: 149473399	Covering
CSF2RA	9	954	ChrY: 1347318	Covering
DBC1	6	534	Chr9: 3170345	Covering
DMP1	9	855	Chr4: 88790089	393 bp upstream
DUSP10	9	835	Chr1: 219977791	366 bp upstream
ETV1	6	534	Chr7: 13995314	Covering
FAT	8	738	Chr4: 187883487	1506 bp upstream

FGFBP1	8	739	Chr4: 15549318	259 bp upstream
FOXO3A	7	634	Chr6: 109091532	340 bp upstream
GAK	4	149	Chr4: 831913	1232 bp upstream
GALNAC4S-6ST	4	359	Chr10: 125841514	1579 bp downstream
GNG4	7	644	Chr1: 233881213	536 bp upstream
HCK	2	159	Chr20: 30104375	658 bp downstream
HIST1H1A	4	455	Chr6: 26126462	446 bp upstream
HLF	4	425	Chr17: 50696793	576 bp upstream
HRH1	8	749	Chr3: 11269000	384 bp upstream
IGF2	8	757	Chr11: 2110104	251 upstream
IGFBP6	6	448	Chr12: 51777179	547 bp upstream
LDOC1	4	468	ChrX: 140099530	580 bp upstream
LTBP1	10	994	Chr2: 33212849	Covering
MSC	3	472	Chr8: 72919760	509 bp upstream
NEFL	4	454	Chr8: 24870630	587 bp upstream
NFKB2	5	469	Chr10: 104144243	Covering
NOTCH1	4	429	Chr9: 138532307	13747 bp downstream
PAQR3	6	559	Chr4: 80080121	540 bp upstream
PARP14	10	946	Chr3: 123900580	Covering
PDPN	9	864	Chr1: 13784260	1422 bp downstream
PLGLB1	11	1040	Chr2: 87102776	317 bp upstream
PPFIA4	9	855	Chr1: 201286646	287 bp upstream
PRKCDBP	13	1229	Chr3: 6298431	Covering
PRKD3	7	641	Chr2: 37398132	406 bp upstream
PRLR	4	363	Chr5: 35266751	417 bp upstream
PTPN3	9	832	Chr9: 111265756	199 bp upstream
RBP1	2	116	Chr3: 140741526	246 bp upstream
RGS12	8	662	Chr4: 3344629	3108 bp upstream
SDHC	7	629	Chr1: 159550339	450 bp upstream
SFRP1	2	129	Chr8: 41286179	Covering
SMURF1	8	749	Chr7: 98471133	4511 bp downstream
SOX30	5	442	Chr5: 157012157	151 bp upstream
SSX1	4	349	ChrX: 47999562	246 bp upstream
SSX4	7	651	ChrX: 48127734	Covering
TES	2	149	Chr7: 115637067	749 bp upstream
TGFBR1	6	554	Chr9: 100906692	521 bp upstream

TP53I3	4	354	Chr2: 24161877	645 bp upstream
TPD52	4	334	Chr8: 81247048	657 bp upstream
TRIM10	4	369	Chr6: 30238327	1659 bp upstream
TULP4	10	933	Chr6: 158653332	347 bp upstream
UCHL1	6	546	Chr4: 40953156	529 bp upstream
WRN	2	136	Chr8: 31010340	Covering

[a]Covering indicates a peak center position within 50 bp up- or downstream of the transcription start site.

III

Relapsed childhood high hyperdiploid acute lymphoblastic leukemia: presence of preleukemic ancestral clones and the secondary nature of microdeletions and RTK-RAS mutations

Josef Davidsson,[1] Kajsa Paulsson,[1,2] David Lindgren,[1] Henrik Lilljebjörn,[1] Tracy Chaplin,[2] Erik Forestier,[3] Mette K. Andersen,[4] Ann Nordgren,[5] Richard Rosenquist,[6] Thoas Fioretos,[1] Bryan D. Young,[2] and Bertil Johansson[1]

[1]Department of Clinical Genetics, University and Regional Laboratories, Lund University Hospital, Lund University, Lund, Sweden; [2]Cancer Research UK Medical Oncology Centre, Barts and the London School of Medicine, Queen Mary College, London, United Kingdom; [3]Department of Clinical Sciences, Pediatrics, University of Umeå, Umeå, Sweden; [4]Department of Clinical Genetics, Rigshospitalet, Copenhagen, Denmark; [5]Department of Molecular Medicine and Surgery, Karolinska Institute, Stockholm, Sweden; and [6]Department of Genetics and Pathology, Uppsala University, Uppsala, Sweden

Although childhood high hyperdiploid acute lymphoblastic leukemia is associated with a favorable outcome, 20% of patients still relapse. It is important to identify these patients already at diagnosis to ensure proper risk-stratification. We have investigated 11 paired diagnostic/relapse samples with SNP array and mutation analyses of *FLT3, KRAS, NRAS*, and *PTPN11* in order to identify changes associated with relapse and to ascertain the genetic evolution patterns. Structural changes, mainly cryptic hemizygous deletions, were significantly more common at relapse ($P<0.05$). No single aberration was linked to relapse, but four deletions, involving *IKZF1*, *PAX5*, *CDKN2A/B* or *AK3*, were recurrent. Based on the genetic relationship between the paired samples, three groups were delineated: 1) identical genetic changes at diagnosis and relapse (18%), 2) clonal evolution with all changes at diagnosis being present at relapse (18%), and 3) clonal evolution with some changes conserved, lost or gained (64%), suggesting the presence of a preleukemic clone. This ancestral clone was characterized by numerical changes only, with structural changes and RTK-RAS mutations being secondary to the high hyperdiploid pattern.

Correspondence: Josef Davidsson, Department of Clinical Genetics, Lund University Hospital, SE-221 85 Lund, Sweden; e-mail: josef.davidsson@med.lu.se; phone: +46 46 173398; fax: +46 46 131061.

Introduction

High hyperdiploid (51-67 chromosomes) B-cell precursor acute lymphoblastic leukemia (ALL) is the largest genetic subgroup of childhood ALL, comprising 25% of all such cases. Cytogenetically, high hyperdiploid ALLs are characterized by a nonrandom gain of chromosomes, in particular X, 4, 6, 10, 14, 17, 18, and 21. In addition, microdeletions of *CDKN2A*, *ETV6*, *IKZF1*, *PAX5*, *RB1*, and *TCF3*, and mutations in the *FLT3, KRAS, NRAS*, and *PTPN11* genes, all involved in the RTK-RAS signaling pathway, are relatively frequent.[1]

In recent larger series of children with high hyperdiploid ALL, the event-free survival and overall survival rates have been approximately 80% and 90%, respectively.[2-4] Thus, despite the generally very good prognosis, a substantial proportion of the cases relapse and some patients succumb to the disease. It is of utmost clinical importance to identify these patients up-front in order to assign them to correct risk groups and hence proper treatment regimens. One approach would be to identify specific relapse-associated genetic features already at the time of diagnosis. It has previously been suggested that high hyperdiploid ALLs with structural chromosome changes are associated with a poor outcome; however, this has not been confirmed in later studies.[3,5-7] In fact, there are no genetic changes, including microdeletions and mutated genes, known to correlate with relapse in high hyperdiploid ALL. In contrast, several investigators have identified "favorable" aberrations, reporting a superior outcome for cases harboring certain trisomies, that is +4, +10,+17, and +18;[2,3,8,9] findings that are now used for risk stratification, at least by the Children's Oncology Group (COG) as regards the triple trisomies +4, +10, and +17.[10]

Another approach to identify genetic changes associated with relapse of high hyperdiploid ALL is to compare the genetic features of paired diagnostic and relapse samples. This would also provide valuable information as regards the temporal order of the various chromosome abnormalities and gene mutations and thus enable identification of "preleukemic" ancestral clones as well as secondary, progression-related changes. Considering the high incidence of high hyperdiploid ALL, surprisingly few patients have been analyzed in this respect. In fact, only two studies, both using SNP array analyses, of paired

diagnostic and relapse samples have been published, comprising a total of eight high hyperdiploid ALLs.[11,12] The two patients analyzed by Yang et al.[11] displayed identical numbers of gains and losses at diagnosis and relapse, providing no evidence for genetic clonal evolution at disease recurrence, whereas Mullighan et al.[12] observed differences in copy number alterations in diagnostic and relapse samples in five of the six cases analyzed. Thus, available data are contradictory as regards genetic heterogeneity/homogeneity between high hyperdiploid ALLs at diagnosis and relapse, and further studies are clearly needed to clarify this biologically and clinically important issue. In the present investigation, we used 500k SNP arrays to analyze paired samples from 11 children with high hyperdiploid ALL, the largest series to date, and also screened for mutations in the RTK-RAS genes *FLT3, KRAS, NRAS*, and *PTPN11*. Apart from a very recent study on the secondary nature of *PTPN11* mutations in high hyperdiploid ALLs,[13] a similar screening has previously not been performed in relapse samples.

Methods

Patient samples

Diagnostic and relapse samples from 11 children with high hyperdiploid ALL were included in the study (Table 1). All patients had been diagnosed and treated according to the NOPHO-ALL 1992 or NOPHO-ALL 2000 protocols at the Departments of Pediatric Oncology and Hematology, Lund University Hospital, Lund, Sweden; Karolinska Institute, Stockholm, Sweden; Linköping University Hospital, Linköping, Sweden; Umeå University Hospital, Umeå, Sweden, Uppsala University Hospital, Uppsala, Sweden; and Rigshospitalet, Copenhagen, Denmark. Cytogenetic analyses of bone marrow samples obtained at the time of diagnosis and relapse were performed using conventional methods. High hyperdiploidy was defined as 51-67 chromosomes.[1] None of the cases had the well-known ALL-associated translocations t(1;19)(q23;p13), t(9;22)(q34;q11), 11q23/*MLL* rearrangements, or t(12;21)(p13;q22). DNA was extracted using standard methods and purified further with the DNeasy Blood & Tissue Kit (Qiagen, Solna, Sweden) and quantified using a Nanodrop

Spectrophotometer (Nanodrop Technologies, Wilmington, DE). The study was reviewed and approved by the research ethics board at Lund University, and informed consent was obtained in accordance with the Declaration of Helsinki.

SNP array analyses

The SNP array analyses were performed using the Affymetrix GeneChip Human 250k *Nsp*, 250k *Sty*, and 10k 2.0 array systems. These arrays cover approximately 510,000 SNPs, with a median physical distance between the SNPs of <2.5 kb. Hybridization and washes were performed as described previously.[14] The Affymetrix GTYPE software was used for analysis of signal intensity and for genotype calling. Apart from copy number analysis, the genotype data were also used to confirm that the diagnostic and relapse samples were derived from the same individual. As an initial step to identify putative genomic imbalances, log2 ratios were segmented with the CBS algorithm[15] using the DNA copy package in R (http://www.bioconductor.org). For further and detailed copy number alteration (CNA) and uni-parental isodisomy (UPD) analyses the dChip[16] and the in-house Genome Orientated Laboratory File (GOLF) software packages were used. When applying the former software for CNA detection, the median smoothing was set to a 21 SNP window and the sample trimming to 20%, with the function "scale copy number mode to 2 copy" being used.

Classification as a CNA was based on visual inspection of the inferred log2 ratio versus the pooled signal intensity of 10 control samples. Ratios +/- ≥0.5 were classified as abnormal, with ratios between +/- 0.5 and 1.0 interpreted as duplications or hemizygous deletions and +/- ≥1.0 classified as amplifications or homozygous deletions. Imbalances representing copy number poly-morphisms, listed in the Database of Genomic Variants (http://projects.tcag.ca/variation/), or somatic immunoglobulin and T cell receptor gene rearrangements were excluded from further analysis.

Mutation analyses of *FLT3*, *KRAS*, *NRAS*, and *PTPN11*

For analyses of codons 835 and 836 in the second tyrosine kinase domain and of internal tandem duplication of exons 14 and 15 in *FLT3*, codons 12, 13, and 61 in *NRAS* and *KRAS* and of exons 3 and 13 in *PTPN11* eight different PCR reactions were performed, as described previously.[17] For each PCR reaction, approximately 100 ng DNA was used as a template. The PCR products were directly sequenced using the McLAB sequencing service (San Francisco, CA) and both the sense and the antisense strands were sequenced and analyzed using the Seqscape software (PE Applied Biosystems, Foster City, CA).

Results

Comparison of paired diagnostic and relapse samples identifies ancestral clones

Based on the genetic relationship between the diagnostic and relapse samples in the 11 patients (Table 1), three distinct groups were identified (Figure 1): 1) **identical** genetic changes at diagnosis and relapse (#1 and 3; 18%), 2) **clonal evolution** with all changes at diagnosis being present at relapse (#6 and 9; 18%), and 3) **ancestral clones** with some changes present at diagnosis but not at relapse (#2, 4, 5, 7, 8, 10, and 11; 64%), suggesting the presence of a preleukemic clone.[12] As seen in Table 1, it can be deduced that the ancestral clones were characterized by numerical changes only, with structural changes and mutated RTK-RAS genes differing between the paired samples.

There were no apparent differences as regards time from diagnosis to relapse among the three groups (4 years, 1-4 years, and 2-6 years, respectively; Table 1).

Patterns of gained chromosomes

The number of gains ranged between 7 and 13 (median 9) among the 11 cases (22 samples), and involved, in decreasing frequency order, chromosomes 6 (100%), 21 (100%), 4 (91%), 17 (91%), 18 (91%), X (86%), 10 (82%), 8 (77%), 14 (45%), 5

(23%), 2 (9%), 7 (9%), 9 (9%), 15 (9%), and 16 (5%). In seven cases (#1, 3-6, 9, and 10), the gains were identical at diagnosis and relapse, whereas they differed by 1-5 trisomies/tetrasomies, involving chromosomes X, 5, 8, 14, 16, and 21, in four (36%) of the 11 patients. There were no major modal shifts in the latter four cases (55→54 in case 2, 55→58 in case 7, 59→58 in case 8, and 57→55 in case 11; Table 1). An acquired UPD was detected in one (9%) of the 11 cases, comprising a UPD8 at diagnosis and relapse in case 8.

Simultaneous trisomies of chromosomes 4, 10, and 17 – the triple trisomies used for stratification of patients into a lower-risk group by the COG[10] – were found in seven (64%) of the 11 patients both at diagnosis and at relapse (Table 1).

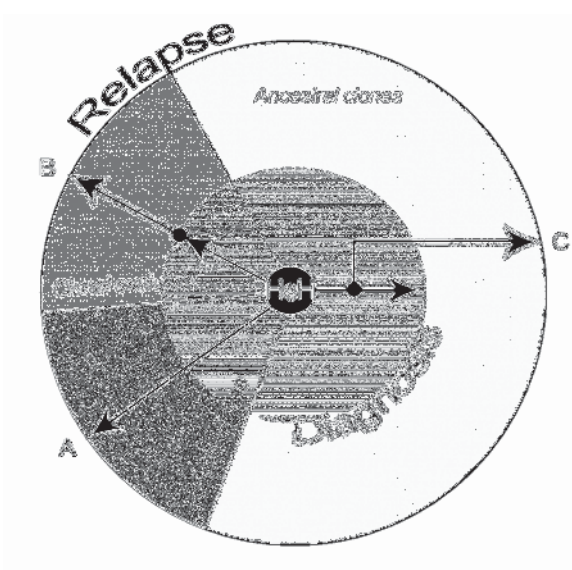


Figure 1. Three genetic groups of relapsed high hyperdiploid ALLs. Based on the genetic relationship between the diagnostic and relapse samples three distinct groups could be delineated. (A) Identical clones (18%) at diagnosis and relapse, (B) Clonal evolution (18%) with all changes at diagnosis being present at relapse, and (C) Ancestral clones (64%), that is some changes were conserved and others lost or gained in the diagnostic and relapse samples, respectively. HeH indicates high hyperdiploidy.

Microdeletions are significantly more common in relapses

The number of structural changes varied between 0 and 6 (median 3) among the 22 samples. A total of 34 different hemizygous deletions, two homozygous deletions, and five copy-number gains (whole chromosomes excluded) were identified. The median size of the 36 hemi- and homozygous deletions was 0.156 Mb (range 0.019-66.579 Mb), with both the homozygous deletions being <1 Mb. The median size of the copy-number gains was 40.127 Mb

(range 0.142-100.895 Mb). Among all 41 structural changes detected, nine (22%) were >10 Mb in size; the remaining 32 (78%) imbalances would be considered cytogenetically cryptic (Table 2).

None of the five copy-number gains involved the same chromosome regions (Table 1). Of the 36 hemi- and homozygous deletions identified, four were overlapping (albeit with different breakpoints) in more than one case and were hence recurrent (Table 2): 7p12.2/*IKZF1* (#4D and 7R), 9p13.2/*PAX5* (#4D and 10D/R), 9p21.3/*CDKN2A/B* (#2D and 4R), and 9p24.1/*AK3* (#8R and 10D) (Figure 2).

None of the homozygous deletions or copy-number gains was shared in paired diagnostic and relapse samples, and apart from a homozygous 9p deletion at diagnosis in case 2, these changes were only found at relapse. Among the 34 hemizygous deletions, nine were seen both at diagnosis and relapse in five patients (#1, 4, 5, 10, and 11). Of the remaining 25 hemizygous deletions, 17 (68%) were found at relapse and 8 (32%) at diagnosis (Tables 1 and 2). Taken together, the mean frequency of structural changes at diagnosis was 1.6 (median 1.0; range 0-4), whereas the corresponding frequency at relapse was 2.9 (median 3.0; range 0-6) ($P<0.05$; Wilcoxon matched pairs signed rank sum test).

Most *FLT3*, *KRAS*, *NRAS*, and *PTPN11* mutations are found at relapses

Five (45%) of the 11 cases harbored a total of six mutated RTK-RAS genes (Table 1): one *FLT3* mutation (#8R), two *KRAS* mutations (#6R and 9R), and three *PTPN11* mutations (#5D, 8D/R, and 10R).

Among the six different mutations, one was detected both at diagnosis and relapse (*PTPN11* in case 8). Of the remaining five mutations, one was found at diagnosis only (#5) and four at relapse only (#6, 8, 9, and 10) (Table 1).

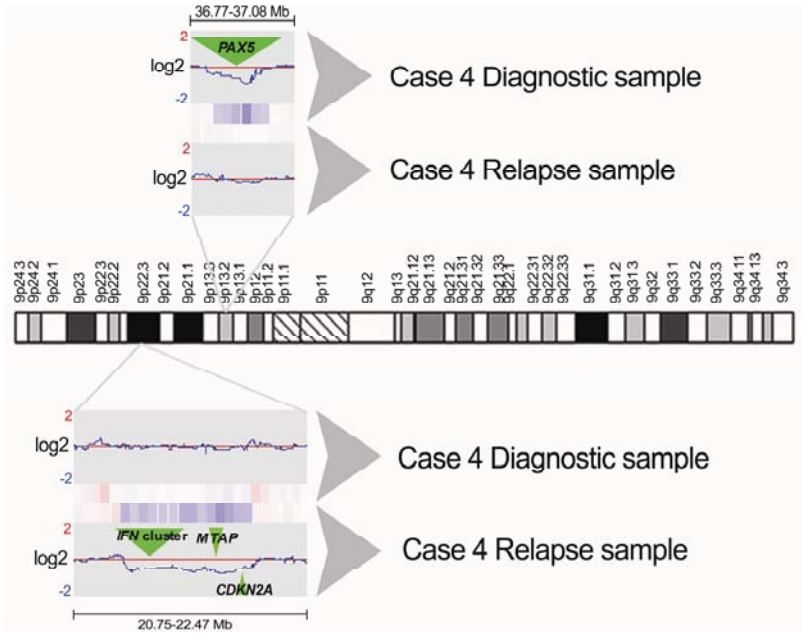


Figure 2. Different 9p deletions in case 4 at diagnosis and at relapse. The SNP array analyses at diagnosis revealed a hemizygous deletion between positions 36.843 Mb and 36.999 Mb at 9p13.2, involving *PAX5*; this deletion was not seen at relapse. I contrast, a hemizygous deletion between positions 21.064 Mb and 22.089 Mb at 9p21.3, involving several genes, such as *CDKN2A*, was identified only at relapse.

Discussion

The salient findings in the present study of paired diagnostic and relapse samples from childhood high hyperdiploid ALLs are 1) that structural changes were significantly more common in relapse samples, 2) that no single, recurrent aberration could be directly linked to relapse and, 3) that there were genetic differences, involving mainly structural abnormalities and mutated RTK-RAS genes, between diagnostic and relapse samples in the vast majority of patients, suggesting the presence of an ancestral high hyperdiploid clone in many of the cases and that structural changes and mutations are secondary to the high hyperdiploid pattern.

Despite the fact that simultaneous occurrence of +4, +10, and +17 in childhood ALL has been associated with low-risk pediatric ALL,[10] we nevertheless found these triple trisomies in the majority (64%) of the patients that later relapsed (Table 1), a

frequency on a par with the incidence of 55% of these triple trisomies in close to 60 high hyperdiploid cases also analyzed by high resolution arrays at the time of diagnosis.[18-21] Thus, it is clear that the triple trisomies are seen in a substantial proportion of relapses of high hyperdiploid ALLs, a finding that may question their favorable prognostic impact.

The frequencies of structural aberrations, mainly cytogenetically cryptic hemizygous deletions, were significantly higher in the relapse samples. This is not unexpected considering that several previous studies, using conventional chromosome banding techniques, have shown increased numbers of structural chromosome abnormalities in ALL, including high hyperdiploid cases, at relapse.[22-27] Using SNP array analysis, Mullighan et al.[12] also observed significantly higher frequencies of deletions in relapse samples. In the present study, most RTK-RAS mutations were detected at relapse. Although not statistically significant, it agrees well with the notion that additional genetic changes are acquired prior to relapse. Whether these are caused by the previous genotoxic therapy or are a consequence of inherent clonal evolution over time is unknown, and may well be impossible to clarify because ALL treatment is solely based on chemotherapy.

There were no structural aberrations or mutated RTK-RAS genes common to all cases (Tables 1 and 2), showing that there is no single aberration directly linked to relapse, at least not identifiable by the methods used herein. However, four structural abnormalities, namely the deletions involving *AK3*, *CDKN2A/B*, *IKZF1*, or *PAX5*, were recurrent in the present series, being found in two cases each (Table 2). Submicroscopic deletions of these genes have previously been identified at diagnosis as well as at relapse of high hyperdiploid ALLs,[1,12,21] but their prognostic impact remains to be clarified, at least in the context of high hyperdiploidy.[28,29] Recently, deletions of *IKZF1* were shown to be a predictor of poor outcome in high-risk pediatric B-cell progenitor ALL.[30] However, high hyperdiploid ALLs are rarely grouped as high-risk.[1] Furthermore, of the four high hyperdiploid cases included in the high-risk group reported by Mullighan et al.[30], none had *IKZF1* deletions, whereas the frequency of such deletions was approximately 10% (4/44) in the validation cohort. Also, among the present 11 patients and the six informative paired samples previously reported[12] three (18%) harbored

IKZF1 deletions, with the deletions being present both at diagnosis and relapse in one case, only at diagnosis in one, and only at relapse in one. Taken together, it seems unlikely that *IKZF1* deletions have a major prognostic impact in high hyperdiploid ALL.

For many years, childhood high hyperdiploid childhood ALLs were thought to be genetically characterized only by the nonrandom tri- and tetrasomies and a handful of well-defined structural chromosome abnormalities. However, with the recent advent of array-based screening methods, revealing a more complex pattern with several and heterogeneous submicroscopic aberrations, this view is now changing.[1] In addition, and as clearly seen in the present series (Tables 1 and 2) and in the study by Mullighan et al.[30], the complexity increases even further when comparing diagnostic and relapse samples. In fact, the genetic relationship between diagnostic and relapse samples can be divided into three distinct groups (Figure 1): identical clones (~20%), unidirectional clonal evolution (~20%), and ancestral clones (~65%). The latter were identifiable by displaying a complex genetic evolution pattern with some changes conserved and others lost or gained and probably correspond to early preleukemic clones from which both the diagnostic and relapse clones emanate (Figure 2).[12]

Although the ancestral clone cannot be directly observed, it is possible to deduce its basic genetic constitution by comparing diagnostic and relapse samples (Tables 1 and 2). It then becomes clear that both structural chromosome changes and mutated RTK-RAS genes are secondary to the high hyperdiploid pattern. Considering that high hyperdiploid cells have been detected already at birth in children who subsequently developed high hyperdiploid ALL,[31,32] the present results indicate that structural chromosome abnormalities and *FLT3, KRAS, NRAS*, and *PTPN11* mutations may represent the secondary changes necessary for overt leukemia.

Acknowledgments

This work was supported by grants from the Swedish Childhood Cancer Foundation, the Swedish Cancer Society, and the Swedish Research Council.

References

1. Paulsson K, Johansson B. High hyperdiploid childhood acute lymphoblastic leukemia. *Genes Chromosomes Cancer* 2009; **48:** 637-660.

2. Heerema NA, Sather HN, Sensel MG, Zhang T, Hutchinson RJ, Nachman JB*, et al.* Prognostic impact of trisomies of chromosomes 10, 17, and 5 among children with acute lymphoblastic leukemia and high hyperdiploidy (> 50 chromosomes). *J Clin Oncol* 2000; **18:** 1876-1887.

3. Moorman AV, Richards SM, Martineau M, Luk Cheung K, Robinson HM, Reza Jalali G*, et al.* Outcome heterogeneity in childhood high-hyperdiploid acute lymphoblastic leukemia. *Blood* 2003; **102:** 2756-2762.

4. Forestier E, Heyman M, Andersen MK, Autio K, Blennow E, Borgström G*, et al.* Outcome of *ETV6/RUNX1*-positive childhood acute lymphoblastic leukaemia in the NOPHO-ALL-1992 protocol: frequent late relapses but good overall survival. *Br J Haematol* 2008; **140:** 665-672.

5. Pui C-H, Raimondi SC, Dodge RK, Rivera GK, Fuchs LA, Abromowitch M*, et al.* Prognostic importance of structural chromosomal abnormalities in children with hyperdiploid (>50 chromosomes) acute lymphoblastic leukemia. *Blood* 1989; **73:** 1963-1967.

6. Raimondi SC, Pui C-H, Hancock ML, Behm FG, Filatov L, Rivera GK. Heterogeneity of hyperdiploid (51-67) childhood acute lymphoblastic leukemia. *Leukemia* 1996; **10:** 213-224.

7. Forestier E, Johansson B, Gustafsson G, Borgström G, Kerndrup G, Johannsson J*, et al.* Prognostic impact of karyotypic findings in childhood acute lymphoblastic leukaemia: a Nordic series comparing two treatment periods. *Br J Haematol* 2000; **110:** 147-153.

8. Harris MB, Shuster JJ, Carroll A, Look AT, Borowitz MJ, Crist WM*, et al.* Trisomy of leukemic cell chromosomes 4 and 10 identifies children with B-progenitor cell acute lymphoblastic

leukemia with a very low risk of treatment failure: a Pediatric Oncology Group study. *Blood* 1992; **79:** 3316-3324.

9. Sutcliffe MJ, Shuster JJ, Sather HN, Camitta BM, Pullen J, Schultz KR*, et al.* High concordance from independent studies by the Children's Cancer Group (CCG) and Pediatric Oncology Group (POG) associating favorable prognosis with combined trisomies 4, 10, and 17 in children with NCI standard-risk B-precursor acute lymphoblastic leukemia: a Children's Oncology Group (COG) initiative. *Leukemia* 2005; **19:** 734-740.

10. Schultz KR, Pullen DJ, Sather HN, Shuster JJ, Devidas M, Borowitz MJ*, et al.* Risk- and response-based classification of childhood B-precursor acute lymphoblastic leukemia: a combined analysis of prognostic markers from the Pediatric Oncology Group (POG) and Children's Cancer Group (CCG). *Blood* 2007; **109:** 926-935.

11. Yang JJ, Bhojwani D, Yang W, Cai X, Stocco G, Crews K*, et al.* Genome-wide copy number profiling reveals molecular evolution from diagnosis to relapse in childhood acute lymphoblastic leukemia. *Blood* 2008; **112:** 4178-4183.

12. Mullighan CG, Phillips LA, Su X, Ma J, Miller CB, Shurtleff SA*, et al.* Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. *Science* 2008; **322:** 1377-1380.

13. Molteni CG, Te Kronnie G, Bicciato S, Villa T, Tartaglia M, Basso G*, et al. PTPN11* mutations in childhood acute lymphoblastic leukemia occur as a secondary event associated with high hyperdiploidy. *Leukemia* 2009; (in press).

14. Paulsson K, Cazier JB, Macdougall F, Stevens J, Stasevich I, Vrcelj N*, et al.* Microdeletions are a general feature of adult and adolescent acute lymphoblastic leukemia: unexpected similarities with pediatric disease. *Proc Natl Acad Sci U S A* 2008; **105:** 6708-6713.

15. Venkatraman ES, Olshen AB. A faster circular binary segmentation algorithm for the analysis of array CGH data. *Bioinformatics* 2007; **23:** 657-663.

16. Lin M, Wei L-J, Sellers WR, Lieberfarb M, Wong WH, Li C. dChipSNP: significance curve and clustering of SNP-array-based loss-of-heterozygosity data. *Bioinformatics* 2004; **20:** 1233-1240.

17. Paulsson K, Horvat A, Strömbeck B, Nilsson F, Heldrup J, Behrendtz M*, et al.* Mutations of *FLT3*, *NRAS*, *KRAS*, and *PTPN11* are frequent and possibly mutually exclusive in high hyperdiploid childhood acute lymphoblastic leukemia. *Genes Chromosomes Cancer* 2007; **47:** 26-33.

18. Strefford JC, Worley H, Barber K, Wright S, Stewart ARM, Robinson HM*, et al.* Genome complexity in acute lymphoblastic leukemia is revealed by array-based comparative genomic hybridization. *Oncogene* 2007; **26:** 4306-4318.

19. Davidsson J, Andersson A, Paulsson K, Heidenblad M, Isaksson M, Borg A*, et al.* Tiling resolution array comparative genomic hybridization, expression and methylation analyses of dup(1q) in Burkitt lymphomas and pediatric high hyperdiploid acute lymphoblastic leukemias reveal clustered near-centromeric breakpoints and overexpression of genes in 1q22-32.3. *Hum Mol Genet* 2007; **16:** 2215-2225.

20. Paulsson K, Heidenblad M, Mörse H, Borg Å, Fioretos T, Johansson B. Identification of cryptic aberrations and characterization of translocation breakpoints using array CGH in high hyperdiploid childhood acute lymphoblastic leukemia. *Leukemia* 2006; **20:** 2002-2007.

21. Mullighan CG, Goorha S, Radtke I, Miller CB, Coustan-Smith E, Dalton JD*, et al.* Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* 2007; **446:** 758-764.

22. Secker-Walker LM, Alimena G, Bloomfield CD, Kaneko Y, Whang-Peng J, Arthur DC*, et al.* Cytogenetic studies of 21 patients with acute lymphoblastic leukemia in relapse. *Cancer Genet Cytogenet* 1989; **40:** 163-169.

23. Shikano T, Ishikawa Y, Ohkawa M, Hatayama Y, Nakadate H, Hatae Y*, et al.* Karyotypic changes from initial diagnosis to relapse in childhood acute leukemia. *Leukemia* 1990; **4:** 419-422.

24. Abshire TC, Buchanan GR, Jackson JF, Shuster JJ, Brock B, Head D*, et al.* Morphologic, immunologic and cytogenetic studies in children with acute lymphoblastic leukemia at diagnosis and relapse: a Pediatric Oncology Group study. *Leukemia* 1992; **6:** 357-362.

25. Heerema NA, Palmer CG, Weetman R, Bertolone S. Cytogenetic analysis in relapsed childhood acute lymphoblastic leukemia. *Leukemia* 1992; **6:** 185-192.

26. Vora AJ, Potter AM, Anderson LM, Lilleyman JS. Frequency and importance of change in blast cell karyotype in relapsing childhood lymphoblastic leukemia. *Pediatr Hematol Oncol* 1994; **11:** 379-386.

27. Chucrallah AE, Stass SA, Huh YO, Albitar M, Kantarjian HM. Adult acute lymphoblastic leukemia at relapse. Cytogenetic, immunophenotypic, and molecular changes. *Cancer* 1995; **76:** 985-991.

28. Sulong S, Moorman AV, Irving JAE, Strefford JC, Konn ZJ, Case MC*, et al.* A comprehensive analysis of the CDKN2A gene in childhood acute lymphoblastic leukemia reveals genomic deletion, copy number neutral loss of heterozygosity, and association with specific cytogenetic subgroups. *Blood* 2009; **113:** 100-107.

29. Den Boer ML, van Slegtenhorst M, De Menezes RX, Cheok MH, Buijs-Gladdines JGCAM, Peters STCM*, et al.* A subtype of childhood acute lymphoblastic leukaemia with poor treatment outcome: a genome-wide classification study. *Lancet Oncol* 2009; **10:** 125-134.

30. Mullighan CG, Su X, Zhang J, Radtke I, Phillips LA, Miller CB*, et al.* Deletion of *IKZF1* and prognosis in acute lymphoblastic leukemia. *N Engl J Med* 2009; **360:** 470-480.

31. Panzer-Grümayer ER, Fasching K, Panzer S, Hettinger K, Schmitt K, Stöckler-Ipsiroglu S*, et al.* Nondisjunction of chromosomes leading to hyperdiploid childhood B-cell precursor acute lymphoblastic leukemia is an early event during leukemogenesis. *Blood* 2002; **100:** 347-349.

32. Maia AT, Tussiwand R, Cazzaniga G, Rebulla P, Colman S, Biondi A*, et al.* Identification of preleukemic precursors of hyperdiploid acute lymphoblastic leukemia in cord blood. *Genes Chromosomes Cancer* 2004; **40:** 38-43.

Table 1. Clinical and genetic features at the time of diagnosis and relapse of childhood high hyperdiploid acute lymphoblastic leukemia

Sample	Sex/ age	Genomic imbalances/UPDs identified by SNP array analyses[a]	*FLT3*, *KRAS*, *NRAS*, and *PTPN11* mutations[a]	Genetic relationship between diagnosis and relapse
1D	F/2	+X,+4,del(4)(p16.1p16.3),del(5)(q35.1q35.1),+6,del(6)(p21.1p22.1),del(11)(q12.1q12.1),+14,+17,+21,+21	No	Identical clones
1R	F/6	+X,+4,del(4)(p16.1p16.3),del(5)(q35.1q35.1),+6,del(6)(p21.1p22.1),del(11)(q12.1q12.1),+14,+17,+21,+21	No	
3D	F/16	+X,+4,+6,+8,+10,+17,+18,+21,+21	No	Identical clones
3R	F/20	+X,+4,+6,+8,+10,+17,+18,+21,+21	No	
6D	F/3	+X,+4,+6,+8,+10,+17,+18,+21,+21	No	Clonal evolution
6R	F/7	+X,+4,**dup(5)(q31.3q35.3)**,+6,+8,**del(9)(q21.1q21.1)**,+10,**del(12)(p12.3p13.3)**,+17,+18,+21,+21	***KRAS*: c.34G>A (p.Gly12Ser)**	
9D	F/2	+4,+6,+8,+10,+15,+17,+18,+21,+21	No	Clonal evolution
9R	F/3	+4,+6,+8,+10,+15,+17,+18,+21,+21	***KRAS*: c.35G>A (p.Gly12Asp)**	
2D	M/5	+X,+2,+4,+6,+8,**del(9)(p21.3p21.3)x1-2**,+10,+18,+21,**+21**	No	Ancestral clone
2R	M/7	+X,+2,+4,**del(4)(p15.3p15.3)**,+6,**del(7)(q33q33)**,+8,**del(9)(q33.1q33.1)**,+10,+18,+21	No	
4D	M/2	+X,+X,+4,+6,**del(7)(p12.2p12.2)**,**del(9)(p13.2p13.2)**,**del(10)(q21.1q21.1)**,+14,+17,+18,**-19**,+21,del(21)(q22.3q22.3)	No	Ancestral clone
4R	M/4	+X,+X,+4,+6,**del(9)(p21.3p21.3)**,+14,+17,+18,**del(20)(q11.2q11.2)**,+21,del(21)(q22.3q22.3)	No	

5D	M/14	+X,+X,del(3)(q23q25.1),+6,+8,+10,+14,+17,+18,+21,+21	***PTPN11*: c.214G>A (p.Ala72Thr)**	Ancestral clone
5R	M/19	+X,+X,del(3)(q23q25.1),+6,+8,+10,**del(10)(q22.1q26.3)**,+14,+17,+18, **dup(20)(q11.1q11.2)**,+21,+21	No	
7D	M/11	+4,+6,+8,+9,+10,**+14**,+17,+18,+21	No	Ancestral clone
7R	M/16	**+X**,**+X**,+4,**+5**,+6,**del(7)(p11.2p22.3)**,**dup(7)(q11.1q36.3)**,+8,+9,+10,**del(11) (p15.1p15.1)**,**del(13)(q14.1q14.3)x1-2**,+17,+18,+21,**+21**	No	
8D	M/5	+X,**+X**,+4,+5,+6,+7,+8,+10,+14,+17,+18,+21,+21 UPD: 8	*PTPN11*: c.215C>T (p.Ala72Val)	Ancestral clone
8R	M/10	+X,+4,+5,+6,+7,+8,**del(9)(p24.1p24.1)**,+10,**del(11)(q24.2q24.2)**,**del(13) (q12.1q12.1)**,+14,+17,+18,+21,+21 UPD: 8	***FLT3:* c.2505-2507delATC (p.Ile836del)** *PTPN11*: c.215C>T (p.Ala72Val)	
10D	F/5	+X,+4,+6,+8,del(9)(p13.2p13.2),**del(9)(p24.1p24.1)**,+10,del(13)(q12.2q12.2), +17,+18,+21,+21	No	Ancestral clone
10R	F/9	+X,+4,+6,+8,del(9)(p13.2p13.2),+10,del(13)(q12.2q12.2),+17,+18,+21,+21	***PTPN11*: c.218C>T (p.Thr73Ile)**	
11D	M/3	+X,**del(1)(q21.1q21.1)**,del(3)(q26.3q26.3),+4,+5,+6,**+8**,+10,**+14**,+17,**del(17) (q11.2q11.2)**,+18,**del(19)(q13.3q13.3)**,+21,**+21**	No	Ancestral clone
11R	M/9	+X,**dup(1)(q11q32.2)**,**del(2)(q23.3q23.3)**,del(3)(q26.3q26.3),+4,+5,+6,+10, **del(11)(q23.1q23.1)**,**+16**,**del(16)(q21q22.1)**,+17,+18,**dup(19)(p13.1p13.1)**, +21	No	

Abbreviations: D, sample from the time of diagnosis; R, sample from the time relapse; F, female; M, male; and UPD, uniparental disomy.

[a]Genomic imbalances and mutations differing between diagnostic and relapse samples are indicated in bold type.

Table 2. Partial chromosomal imbalances in paired diagnostic and relapsed samples from childhood high hyperdiploid acute lymphoblastic leukemias

Case	Chr	Region	Type/ copies[a]	Position (Mb)[b]	Size (Mb)	Genes[c]
1D/R	4	4p16.1-16.3	He L/2	0-9.820	9.820	Several
1D/R	5	5q35.1	He L/1	169.244-169.539	0.295	*DOCK2/FOXI1/LOC345430*
1D/R	6	6p21.1-22.1	He L/2	29.683-44.190	14.507	Several
1D/R	11	11q12.1	He L/1	56.798-57.135	0.337	*TNKS1BP1/SSRP1/P2RX3/PRG3/PRG2/SLC43A3/RTN4RL2/SLC43A1/TIMM10/SMTNL1/ UBE2L6/LOC390183/SERPING1*
2R	4	4p15.3	He L/2	20.146-20.245	0.099	*SLIT2*
2R	7	7q33	He L/1	134.009-134.144	0.135	*CALD1/LOC340351*
2D	9	9p21.3	He L/1	20.675-22.583	1.908	*KIAA1797/LOC401494/IFNB1/IFNW1/IFNA21/LOC392289/IFNA4/LOC392290/IFNA7/*
		9p21.3	Ho L/0	21.775-22.583	0.808	*IFNA10/G13P1/LOC392291/IFNA16/IFNA17/LOC392292/IFNA14/IFNAP22/IFNA5/ KLHL9/IFNA6/IFNA13/IFNA2/IFNA8/LOC402357/IFNA1/LOC402358/IFNE1/ LOC441388/LOC402359/MTAP/**CDKN2A/CDKN2B/LOC401495/DMRTA1***
2R	9	9q33.1	He L/1	115.116-115.206	0.090	*DEC1*
4D	7	7p12.2	He L/1	50.143-50.320	0.177	*IKZF1*
4R	9	9p21.3	He L/1	21.064-22.089	1.025	*IFNB1/IFNW1/IFNA21/LOC392289/IFNA4/LOC392290/IFNA7/IFNA10/G13P1/ LOC392291/IFNA16/IFNA17/LOC392292/IFNA14/IFNAP22/IFNA5/KLHL9/IFNA6/ IFNA13/IFNA2/IFNA8/LOC402357/IFNA1/LOC402358/IFNE1/LOC441388/LOC402359/ MTAP/CDKN2A/CDKN2B/LOC401495*
4D	9	9p13.2	He L/1	36.843-36.999	0.156	*PAX5*
4D	10	10q21.1	He L/1	53.702-53.783	0.081	*PRKG1/DKK1*
4R	20	20q11.2	He L/1	32.635-32.805	0.170	*CDC91L1/TP53INP2/NCOA6*

4D/R	21	21q22.3	He L/2	42.059-46.885	4.826	Several
5D/R	3	3q23-25.1	He L/1	141.219-153.246	12.027	Several
5R	10	10q22.1-26.3	He L/2	68.743-135.322	66.579	Several
5R	20	20q11.1-11.2	G/3	28.120-36.379	8.259	Several
6R	5	5q31.3-35.3	G/3	140.498-180.625	40.127	Several
6R	9	9q21.1	He L/1	70.605-70.656	0.051	*TRMP3*
6R	12	12p12.3-13.3	He L/1	0-18.253	18.253	Several
7R	7	7p11.2-22.3	He L/1	0-57.731	57.731	Several (including *IKZF1*)
7R	7	7q11.1-36.3	G/3	57.731-158.626	100.895	Several
7R	11	11p15.1	He L/1	18.906-18.939	0.033	*LOC390098/LOC390099/MRGX1/LOC390101*
7R	13	13q14.1-14.3	He L/1	45.175-51.859	6.684	Several
		13q14.2	Ho L/0	47.809-47.973	0.164	***RB1/P2RY5/RCBTB2***
8R	9	9p24.1	He L/1	4.694-4.786	0.092	*CDC37L1/AK3/LOC392282/RCL1*
8R	11	11q24.2	He L/1	126.201-126.258	0.057	*KIRREL3*
8R	13	13q12.1	He L/1	21.127-21.211	0.084	*FGF9*
10D	9	9p24.1	He L/1	4.694-4.784	0.090	*CDC37L1/AK3/LOC392282/RCL1*
10D/R	9	9p13.2	He L/1	36.875-37.002	0.127	*PAX5*
10D/R	13	13q12.2	He L/1	27.529-27.684	0.155	*FLT3/PAN3*
11D	1	1q21.1	He L/1	142.379-142.504	0.125	*PDE4DIP*
11R	1	1q11-32.2	G/3	141.484-206.994	65.510	Several
11R	2	2q23.3	He L/1	152.918-153.025	0.107	*FMNL2*
11D/R	3	3q26.3	He L/1	178.381-178.583	0.202	*TBL1XR1*
11R	11	11q23.1	He L/1	112.332-112.351	0.019	*NCAM1*
11R	16	16q21-22.1	He L/2	56.856-67.106	10.250	Several
11D	17	17q11.2	He L/2	26.171-26.693	0.522	*CRLF3/FLJ12735/FLJ22729/CENTA2/RNF135/LOC400590/NF1/OMG/EVI2B/EVI2A*

11R	19	19p13.1	G/3	15.974-16.116	0.142	*LOC126536/FLJ25328/OR1AB1P/TPM4/RAB8A/HSH2D*
11D	19	19q13.3	He L/1	53.246-53.344	0.098	*PLA2G4C/LIG1*

Abbreviations: D, indicates sample from the time of diagnosis; R, sample from the time relapse; Chr, chromosome; Mb, megabase.

[a]Type: hemizygous loss (He L), homozygous loss (Ho L) or copy-number gain (G). Copy: number of alleles present in the region involved (note that 2 copies represent a loss if the chromosome was trisomic).

[b]The SNP array platform is based on NCBI Build 35.1.

[c]Genes only listed if the region involved was <2 Mb, in which case all known genes are listed from pter to qter. Bold type indicates homozygously deleted genes.