

Università degli Studi di Milano Bicocca

Facoltà di Medicina e Chirurgia

Dottorato in Ematologia Sperimentale, XXV ciclo



***RECURRENT SETBP1 MUTATIONS IN
ATYPICAL CHRONIC MYELOID LEUKEMIA***

Dr. Simona Valletta

Coordinatore: Prof. Carlo Gambacorti-Passerini

Tutor: Prof. Carlo Gambacorti-Passerini

Dr. Rocco Piazza

Anno accademico 2012-2013

ABSTRACT	3
Introduction	5
Atypical Chronic myeloid leukemia	6
WHO classification of tumors of hematopoietic and lymphoid tissue (2008)	7
Principles of the WHO classification.....	8
The WHO classification of hematopoietic and lymphoid tumors: general features	9
The WHO classification of myeloid neoplasms and acute leukemia, general features	13
Revised WHO classification of myeloid neoplasms and acute leukemia	15
Myeloproliferative neoplasms (MPN)	15
Myeloid and Lymphoid neoplasms with eosinophilia and abnormalities in PDGFRA, PDGFRB and FGFR1 genes	17
Myelodysplastic/myeloproliferative neoplasms (MDS/MPN).....	18
Myelodysplastic syndromes (MDS)	19
Acute myeloid leukemia (AML) and related precursor neoplasms.....	20
Acute leukemias of ambiguous lineage.....	23
Precursor lymphoid neoplasm: B lymphoblastic leukemia/lymphoma and T lymphoblastic leukemia/lymphoma	24
Myeloproliferative neoplasms: Chronic Myeloid Leukemia (CML)	26
Chronic Myeloid Leukemia	26
Biology of Chronic Myeloid Leukemia	27
The physiologic function of the translocation partners	29
Molecular anatomy of the BCR-ABL translocation.....	32
Activated signaling pathways and biologic properties of BCR-ABL positive cells	35
Altered adhesion properties	35
Activation of mitogenic signaling.....	36
CML treatment	43
Atypical Chronic Myeloid Leukemia (aCML)	52
Guidelines for distinguishing chronic myelomonocytic leukemia and atypical chronic myeloid leukemia.....	52
Atypical Chronic Myeloid Leukemia (aCML).....	54
Aim	57
Material and methods	60

Results	80
<i>Exome sequencing of aCML samples</i>	81
<i>Recurrent SETBP1 mutations</i>	95
<i>RNA-seq analysis</i>	103
<i>Clinical course of aCML cases according to SETBP1 mutation status</i>	105
<i>Biological effects of the SETBP1 p.Gly870Ser alteration</i>	107
<i>Intracellular localization of SETBP1</i>	114
Discussion	117
Bibliography	124

ABSTRACT

Atypical chronic myeloid leukemia (aCML) is a heterogeneous disorder belonging to the group of myelodysplastic/myeloproliferative (MDS/MPN) syndromes. aCML shares clinical and laboratory features with Chronic Myeloid Leukemia (CML), but it lacks the Philadelphia chromosome and the resulting *BCR-ABL* fusion gene. This crucial difference with CML points to a different pathogenetic process. Because no specific recurrent genomic or karyotypic abnormalities have been identified in aCML, and the molecular pathogenesis of this disease has remained elusive with a dismal outcome, we performed exome-sequencing of eight aCML patients, in order to identify new possible recurrent mutations. The presence of an identical mutation not previously involved in cancer in two different aCML cases altering *SETBP1* gene, prompted us to resequence this particular gene in samples from additional subjects with aCML or other hematological malignancies and in cell lines representative of the most common human solid cancers. *SETBP1* mutations were identified only in aCML and in the closely related disorders and represents the first gene shown to be enriched and recurrently mutated in aCML. Most *SETBP1* mutations were located between codons 858 and 871 causing abrogation of a degron binding site for E3-ubiquitin β -TrCP1 and protection from proteasomal degradation. This causes accumulation of SETBP1 and SET protein, decreased PP2A phosphatase activity and higher proliferation rates. Individuals with *SETBP1* mutations had higher white blood cell counts and worse prognosis, indicating *SETBP1* as a possible valuable diagnostic tool in the differential diagnosis of MDS/MPN syndromes and their prognosis.

This study increases the knowledge of the mechanisms by which malignancy arises and will have important consequences for the diagnosis, prognosis and treatment of aCML and diseases associated with *SETBP1* alterations.

Introduction

Atypical Chronic myeloid leukemia

Atypical chronic myeloid leukemia (aCML) is a rare hematopoietic stem cell disorder, which shows both myeloproliferative as well as myelodysplastic features (1). In the past, the term aCML has been used inconsistently for various Ph chromosome negative myeloproliferative neoplasms closely mimicking chronic myelogenous leukemia (CML), and reports on aCML included patients with a broad range of clinical and laboratory features (2). The situation changed in 2001 with the introduction of World Health Organization (WHO) classification system which provided, for the first time, a set of precise criteria to be used for a diagnosis of aCML. The World Health Organization (WHO) classification of myeloid and lymphoid neoplasms utilizes morphology, immunophenotype, genetics and clinical features to define disease entities of clinical significance. It is a consensus classification in which a number of experts have agreed on the classification and diagnostic criteria (3). In addition to the absence of *BCR/ABL* fusion gene, these criteria introduced by WHO classification in order to identify aCML include the presence of persistent unexplained leukocytosis with severe dysgranulopoiesis, no or minimal absolute monocytosis (monocytes <10% of WBC) and no or minimal basophilia (basophils <2%), presence of neutrophil precursors ($\geq 10\%$), and a peripheral blood and bone marrow blast count of less than 20% and absence of *PDGFRA* and *PDGFRB* rearrangements (4). Because of the presence of dysplasia in what would be otherwise considered a myeloproliferative neoplasm, aCML belongs to the myelodysplastic/myeloproliferative neoplasms (MDS/MPN) category of the WHO system. aCML as defined by these criteria is a rare disorder with a relatively poor prognosis, the median survival ranging from 14 to 25 months, and in which transformation to acute

leukemia occurs in 25-40% of patients (1),(5),(6). The molecular pathogenesis of aCML is unknown, and no recurrent cytogenetic abnormalities have been identified.

WHO classification of tumors of hematopoietic and lymphoid tissue (2008)

An ideal classification scheme of hematopoietic malignancies should include diseases that are clinically significant, clearly defined, mutually exclusive of each other, and that can be diagnosed using currently available technology and information. In addition, there should be general consensus and acceptance of the classification for it to be useful for daily clinical practice as well as for scientific investigations. Lastly, the classification should be flexible and changeable as new information accumulates (3). In 2001, the World Health Organization (WHO), in collaboration with the Society for Hematopathology and the European Association of Hematopathology, attempted to satisfy these requirements and published a classification of Tumors of the Hematopoietic and Lymphoid Tissues as part of the 3rd edition of the series, *WHO Classification of Tumors* (4). In 2008, the classification was updated and published as part of the 4th edition of the WHO monograph series (7), in order to incorporate new scientific and clinical information that has accumulated since the previous edition.

Principles of the WHO classification

The concepts that underlie the WHO classification were derived from numerous published clinical and scientific studies and from the experience of more than 100 pathologists, clinicians, and scientists from around the world who collaborated to develop this consensus classification. A basic principle of the WHO system is to classify hematopoietic and lymphoid neoplasms considering not only morphologic findings but also all available information, including genetic, immunophenotypic, biologic, and clinical features to define specific disease entities (8). The relative contribution of each of these parameters to the final diagnosis varies depending on the disease entity. For some neoplasms, in fact, morphology alone may be sufficient for classification, but for others, knowledge of genetic lesions is necessary for the final diagnosis and classification, and often for the treatment as well (3).

For the lymphoid neoplasms, the WHO classification provides refinement of the entities defined in the Revised European-American Lymphoma (REAL) Classification; instead, to classify myeloid neoplasms, the WHO system includes many of the criteria of the French-American-British (FAB) Cooperative Group classifications of acute myeloid leukemia (AML) and myelodysplastic syndromes (MDS) as well as guidelines of the Polycythemia Vera Study Group (PVSG) for the chronic myeloproliferative diseases (CMPDs), but there are also significant differences (8).

The WHO classification of hematopoietic and lymphoid tumors: general features

The complete WHO classification is listed in Table 1.

CLASSIFICATION OF HEMATOPOIETIC AND LYMPHOID NEOPLASMS
<p>Myeloproliferative neoplasms Chronic myelogenous leukemia, <i>BCR-ABL1</i> positive Chronic neutrophilic leukemia Polycythemia vera Primary myelofibrosis Essential thrombocythaemia Chronic eosinophilic leukemia, NOS</p> <p>Mastocytosis Cutaneous mastocytosis Systemic mastocytosis Mast cell leukemia Mast cell sarcoma Extracutaneous mastocytoma Myeloproliferative neoplasm, unclassifiable</p> <p>Myeloid and lymphoid neoplasms with eosinophilia and abnormalities of <i>PDGFRA</i>, <i>PDGFRB</i> or <i>FGFR1</i> Myeloid and lymphoid neoplasms with <i>PDGFRA</i> rearrangement Myeloid neoplasms with <i>PDGFRB</i> rearrangement Myeloid and lymphoid neoplasms with <i>FGFR1</i> abnormalities</p> <p>Myelodysplastic/myeloproliferative neoplasms Chronic myelomonocytic leukemia Atypical chronic myeloid leukemia, <i>BCR-ABL1</i> negative Juvenile myelomonocytic leukemia Myelodysplastic/myeloproliferative neoplasm, unclassifiable Refractory anemia with ring sideroblasts associated with marked thrombocytosis</p> <p>Myelodysplastic syndromes Refractory cytopenia with unilineage dysplasia Refractory anemia Refractory neutropenia Refractory thrombocytopenia Refractory anemia with ring sideroblasts Refractory cytopenia with multilineage dysplasia Refractory anemia with excess blasts Myelodysplastic syndrome associated with isolated del(5q) Myelodysplastic syndrome, unclassifiable Childhood myelodysplastic syndrome Refractory cytopenia of childhood</p> <p>Acute myeloid leukemia (AML) and related precursor neoplasms AML with recurrent genetic abnormalities AML with t(8;21)(q22;q22), <i>RUNX1-RUNX1T1</i> AML with inv(16)(p13.1q22) or t(16;16)(p13.1;p22); <i>CBFB-MYH11</i> Acute promyelocytic leukemia with t(15;17)(q22;q12); <i>PML-RARA</i> AML with t(9;11)(p22;q23) <i>MLLT3-MLL</i> AML with inv(3)(q21q26.2) or t(3.3)(q21;q26.2); <i>RPN1-EVI1</i> AML (megakaryoblastic) with t(1:22)(p13;q13); <i>RBM15-MKL1</i> AML with mutated <i>NPM1</i> AML with mutated <i>CEBPA</i></p> <p>AML with myelodysplasia-related changes Therapy-related myeloid neoplasms</p>

Acute myeloid leukemia, NOS

AML with minimal differentiation
AML without maturation
AML with maturation
Acute myelomonocytic leukemia
Acute monoblastic and monocytic leukemia
Acute erythroid leukemia
Acute megakaryoblastic leukemia
Acute basophilic leukemia
Acute panmyelosis with myelofibrosis

Myeloid sarcoma**Myeloid proliferation related to Down syndrome**

Transient abnormal myelopoiesis
Myeloid leukemia associated with Down syndrome

Blastic plasmacytoid dendritic cell neoplasm**Acute leukemias of ambiguous lineage**

Acute undifferentiated leukemia
Mixed phenotype acute leukemia with t(9;22)(q34;q11.2);BCR-ABL1
Mixed phenotype acute leukemia with t(v;11q23);MLL rearranged
Mixed phenotype acute leukemia, B/myeloid, NOS
Mixed phenotype acute leukemia, T/myeloid, NOS
Natural killer (NK) cell lymphoblastic leukemia/lymphoma

Precursor lymphoid neoplasms

B lymphoblastic leukemia/lymphoma
B lymphoblastic leukemia/lymphoma, NOS
B lymphoblastic leukemia/lymphoma with recurrent genetic abnormalities
B lymphoblastic leukemia/lymphoma with t(9;22)(q34;q11.2); BCR-ABL1
B lymphoblastic leukemia/lymphoma with t(v;11q23);MLL rearranged
B lymphoblastic leukemia/lymphoma with t(12;21)(p13;q22) TEL-AML1 (ETV6-RUNX1)
B lymphoblastic leukemia/lymphoma with hyperdiploidy
B lymphoblastic leukemia/lymphoma with hypodiploidy (hypodiploid ALL)
B lymphoblastic leukemia/lymphoma with t(5;14)(q31;q32) IL3-IGH
B lymphoblastic leukemia/lymphoma with t(1;19)(q23;p13.3);E2A-PAX1
T lymphoblastic leukemia/lymphoma

Mature B-cell neoplasms

Chronic lymphocytic leukemia/small lymphocytic lymphoma
B cell prolymphocytic leukemia
Splenic B-cell marginal zone lymphoma
Hairy cell leukemia
Splenic B-cell lymphoma/leukemia, unclassifiable
Splenic diffuse red pulp small B-cell lymphoma
Hairy cell leukemias –Variant
Lymphoplasmacytic lymphoma
Waldenstrom macroglobulinemia
Heavy chain diseases
Alpha heavy chain disease
Gamma heavy chain disease
Mu heavy chain disease
Plasma cell myeloma
Solitary plasmacytoma of bone
Extrasosseous plasmacytoma
Extranodal marginal zone lymphoma of mucosa-associated lymphoid tissue (MALT lymphoma)
Nodal marginal zone lymphoma
Pediatric nodal marginal zone lymphoma
Follicular lymphoma
Pediatric follicular lymphoma
Primary cutaneous follicle centre lymphoma
Mantle cell lymphoma
Diffuse large B-cell lymphoma (DLBCL), NOS
T-cell/histiocyte rich large B-cell lymphoma
Primary DLBCL of the CNS
Primary cutaneous DLBCL, leg-type
EBV positive DLBCL of the elderly
DLBCL associated with chronic inflammation
Lymphomatoid granulomatosis
Primary mediastinal (thymic) large B-cell lymphoma
Intravascular large B-cell lymphoma

ALK positive large B-cell lymphoma
 Plasmablastic lymphoma
 Large B-cell lymphoma arising in HHV8-associated multicentric Castleman disease
 Primary effusion lymphoma
 Burkitt lymphoma
 B-cell lymphoma, unclassifiable, with features intermediate between diffuse large B-cell lymphoma and Burkitt lymphoma
 B-cell lymphoma, unclassifiable, with features intermediate between diffuse large B-cell lymphoma and classical Hodgkin lymphoma
Mature T-cell and NK-cell neoplasms
 T-cell prolymphocytic leukemia
 T-cell large granular lymphocytic leukemia
 Chronic lymphoproliferative disorder of NK-cells
 Aggressive NK-cell leukemia
 Systemic EBV positive T-cell lymphoproliferative disease of childhood
 Hydroa vacciniforme-like lymphoma
 Adult T-cell leukemia/lymphoma
 Extranodal NK/T cell lymphoma, nasal type
 Enteropathy-associated T-cell lymphoma
 Hepatosplenic T-cell lymphoma
 Subcutaneous panniculitis-like T-cell lymphoma
 Mycosis fungoides
 Sezary syndrome
Primary cutaneous CD30 positive T-cell lymphoproliferative disorders
 Lymphomatoid papulosis
 Primary cutaneous anaplastic large cell lymphoma
 Primary cutaneous gamma-delta T- cell lymphoma
 Primary cutaneous CD8 positive aggressive epidermotropic cytotoxic T-cell lymphoma
 Primary cutaneous CD4 positive small/medium T-cell lymphoma
 Peripheral T-cell lymphoma, NOS
 Angioimmunoblastic T-cell lymphoma
 Anaplastic large cell lymphoma, ALK positive
 Anaplastic large cell lymphoma, ALK negative
Hodgkin lymphoma
 Nodular lymphocyte predominant Hodgkin lymphoma
 Classical Hodgkin lymphoma
 Nodular sclerosis classical Hodgkin lymphoma
 Mixed cellularity classical Hodgkin lymphoma
 Lymphocyte-depleted classical Hodgkin lymphoma
Histiocytic and dendritic cell neoplasms
 Histiocytic sarcoma
 Langerhans cell histiocytosis
 Langerhans cell sarcoma
 Interdigitating dendritic cell sarcoma
 Follicular dendritic cell sarcoma
 Fibroblastic reticular cell tumor
 Indeterminate dendritic cell tumor
 Disseminated Juvenile xanthogranuloma
Post-transplant lymphoproliferative disorders (PTLD)
 Early lesions
 Plasmacytic hyperplasia
 Infectious mononucleosis-like PTLD
 Polymorphic PTLD
 Monomorphic PTLD (B-and T/NK cell types)
 Classical Hodgkin lymphoma type PTLD

Table 1. WHO classification of hematopoietic and lymphoid neoplasms (J W Vardiman. The World Health Organization (WHO) classification of tumors of the hematopoietic and lymphoid tissues: An overview with emphasis on the myeloid neoplasms. Chemico-Biological Interactions 2010; 184:16-20)

Table 1 shows that the hematopoietic neoplasms are stratified broadly according to the lineage of the neoplastic cells, in other words myeloid, lymphoid, histiocytic/dendritic, or ambiguous lineage. The latter category consists of precursor cells neoplasms, cells that lack any specific lineage-associated markers and can be identified like “undifferentiated”, or that express antigens of more than one lineage, and thus appear to have a mixed lineage phenotype (7). Neoplasms comprised of precursor cells (acute myeloid leukemia, lymphoblastic leukemia/lymphoma, blastic plasmacytoid dendritic cell neoplasm, and acute leukemia of ambiguous lineage) are considered separately from those comprised of more mature cells (myeloproliferative neoplasms, myelodysplastic/myeloproliferative neoplasms, myelodysplastic syndromes, mature B-cell and T/NK-cell lymphoma, Hodgkin lymphoma and histiocytic/dendritic cell neoplasms) (3). For the mature lymphoid neoplasms, an additional sub-classification is based on the stage of differentiation as compared to a postulated normal counterpart (e.g., mantle cell lymphoma, follicular lymphoma), on morphology (e.g., diffuse large B cell lymphoma), on clinical presentations or the clinical setting (e.g., diffuse large B cell lymphoma associated with chronic inflammation), or more commonly, on the combination of morphologic, immunophenotypic and/or genetic parameters that together allow a specific disease entity to be defined (e.g., Anaplastic large cell lymphoma, ALK positive).

The revised classification shows significant differences from the previous edition, as shown in table 1. Among these changes is the recognition that for some lymphoid neoplasms, the age of the patient or the location of the neoplasm may be so closely tied to the biology of the tumor that this information is included in the nomenclature. Examples of age-related

neoplasms include pediatric follicular lymphoma, which usually presents with localized disease, a high histologic grade, typically no *BCL2-IGH* rearrangement, and a good prognosis (9). In addition, the new classification includes the location of some tumors that has an important impact on the biology of the neoplasm, denoting it like a specific entity. An example is diffuse large B cell lymphoma of the CNS, which is characterized by a distinct gene expression signature, and primary cutaneous diffuse large B cell lymphoma, leg-type, which usually has a more aggressive course than other primary cutaneous large B cell lymphomas (3). However, it is important to remember that although the location of the tumor may relate strongly to the biology of some hematopoietic neoplasms, other disease entities may present as leukemia in the blood or as a tumor in a lymph node or other organ, and the type of presentation is insignificant for its classification.

The WHO classification of myeloid neoplasms and acute leukemia, general features

The complete WHO classification is listed in Table 1 (3).

In the WHO classification, the term “myeloid” includes all cells belonging to the granulocytic (neutrophil, eosinophil, basophil), monocytic/macrophage, erythroid, megakaryocytic and mast cell lineages. The WHO criteria for myeloid neoplasms apply to initial diagnostic peripheral blood (PB) and bone marrow (BM) specimens obtained prior to any definitive therapy for a suspected hematologic neoplasm (10). All available information, including morphologic, cytochemical, and/or immunophenotypic characteristics, are used to establish the lineage of the neoplastic cells and to assess their degree of maturation. In particular, the

blast percentage represents a practical instrument that allows to categorize myeloid neoplasms and to study their progression. Blast percentages should be derived, if possible, using PB and cellular BM aspirate smears stained with Wright-Giemsa, or May-Grünwald Giemsa. The blast percentage should be correlated with an estimate of the blast count from the marrow biopsy section. The definition of blasts harks back to the parameters proposed by the International Working Group on Morphology of Myelodysplastic Syndrome. Although assessment of the number of cells that express the antigen CD34 provides valuable data for diagnostic and prognostic purposes, the percentage of CD34+ cells should not be considered a substitute for a blast count from the smears or an estimate from the bone marrow biopsy (8). In fact, CD34+ hematopoietic cells usually are blasts, but not all blasts express CD34, so the flow cytometry analysis can give misleading results. In the WHO classification, myeloid neoplasms are stratified into those comprised mainly of blasts with minimal if any maturation [acute myeloid leukemia (AML)], and those in which there is maturation, either effective maturation [myeloproliferative neoplasms (MPN)], ineffective maturation with dysplastic features [myelodysplastic syndromes (MDS)] or both ineffective and effective maturation [myelodysplastic/myeloproliferative neoplasms (MDS/MPN)] in the myeloid lineages (3). The 3rd edition of WHO classification included, for the first time in any widely used classification, genetic information as criteria not only for the diagnosis of CML, but for some subtypes of AML as well (3). AML, MDS, and MPN categories include specific genetic subcategories; thus, a complete cytogenetic analysis of BM cells is essential during initial evaluation for establishing a baseline karyotype and must be performed, if possible. Additional genetic studies should be guided by the results of the initial karyotype and by the

diagnosis suspected based on the clinical, morphologic, and immunophenotypic studies (10). Reverse transcriptase-polymerase chain reaction (RT-PCR) and/or fluorescence in situ hybridization (FISH) techniques can be important to detect recurrent genetic abnormalities in the myeloid neoplasms not detected by routine karyotyping, such as the *FIP1L1-PDGFR*A fusion in some myeloid neoplasms associated with eosinophilia. In addition, gene mutations are increasingly being recognized as important diagnostic and prognostic markers in myeloid neoplasms (10), like for example mutations of *JAK2* in MPN, *NRAS* in MDS/MPN and *NPM1* in AML. Despite the important role of genetic abnormalities into the myeloid categories as diagnostic criteria, it is important to say that a multi-parameter approach is still required, to have a precise diagnosis. In fact, for example, detection of monosomy 5 in a myeloid neoplasm does not identify the lesion as MDS or as AML, rather a carefully performed blast count on the bone marrow samples is required for that distinction (3). In some cases, like therapy-related MDS and AML (tMDS/tAML), can be important the clinical setting in which the myeloid neoplasm occurs for the classification.

Revised WHO classification of myeloid neoplasms and acute leukemia

Myeloproliferative neoplasms (MPN)

The nomenclature for the myeloproliferative entities is changed from “myeloproliferative disorders” (MPDs) to “myeloproliferative neoplasms” (MPNs) considering their neoplastic

nature. The revisions in the criteria for the WHO classification of MPN have been influenced by 2 factors:

- 1) the discovery of genetic abnormalities that can be used as diagnostic markers in *BCR/ABL* negative MPN;
- 2) better characterization of histologic features that aid in the identification of MPN subtypes (10).

The identification of clonal rearrangements or mutations of genes that encode surface or cytoplasmatic protein tyrosine kinases (PTKs), such as *BCR/ABL* and *JAK2*, and of myeloid and lymphoid diseases concomitant with *PDGFRA*, *PDGFRB* and *FGFR1* rearrangements has been important to define new disease groups and to change the diagnostic criteria (11). The most commonly recognized mutation in *BCR/ABL* negative MPN is *JAK2* V617F (12). This particular mutation is present in more than 90% of patients with polycythemia vera (PV) and in 50% of patients with primary myelofibrosis (PMF) or essential thrombocythemia (ET). For this reason, the WHO diagnostic algorithms for PV, ET and PMF at the present time require mutational status of *JAK2* gene and if absent, of other correlated genes linked to clonality of the proliferative process. Additional parameters are required to ensure accurate diagnosis and subclassification regardless of whether a mutation is or is not found. Clinical and laboratory findings form the backbone of these diagnostic criteria, but histopathologic findings characteristic of each MPN subtype have been included in the criteria as well (10).

In addition, for the diagnosis of ET, the new revised criteria include the platelet threshold, that has been lowered from $600 \times 10^9/L$ to $450 \times 10^9/L$. An additional change is that some cases that previously met the criteria for Chronic Eosinophilic leukemia (CEL) now are included into

the group of myeloid and lymphoid neoplasms with eosinophilia and abnormalities in *PDGFRA*, *PDGFRB* and *FGFR1* genes. Cases of myeloid neoplasms with eosinophilia that lack these rearrangements are classified as CEL category, not otherwise specified (NOS).

The appreciation of the role abnormal PTKs play in the pathogenesis of the MPNs argued for the inclusion of other chronic myeloid proliferations related to constitutively activated PTKs under the MPN category. Thus, systemic mastocytosis, which has many myeloproliferative features and is almost always associated with the *KIT* D816V mutation, has been added to the MPN group (10).

Myeloid and Lymphoid neoplasms with eosinophilia and abnormalities in PDGFRA, PDGFRB and FGFR1 genes

They constitute a new group in the 2008 WHO classification. These rare neoplasms are defined as clonal diseases caused by abnormalities in genes coding for alpha or beta chains in receptors with tyrosine kinase activity: platelet-derived growth factor receptor (PDGFR) or fibroblast growth factor receptor (FGFR1) (11). The importance of this new group is correlated to its response to tyrosine kinase inhibitors, in particular, imatinib. The most frequent rearrangement of *PDGFRA* gene is the *FIP1L1-PDGFR* fusion, due to the deletion of chromosome region 4q12, detected only using molecular techniques. Karyotyping or fluorescence *in situ* hybridization (FISH) allow the identification of *PDGFRB* (5q33) and *FGFR1* (8p11) genes rearrangements.

Myelodysplastic/myeloproliferative neoplasms (MDS/MPN)

These diseases include clonal myeloid neoplasms that present clinical and laboratory features of both myelodysplasia and myeloproliferation (11). This subgroup includes chronic myelomonocytic leukemia (CMML), atypical chronic myeloid leukemia (aCML), juvenile myelomonocytic leukemia (JMML), and a provisional entity within the MDS/MPN unclassifiable group, refractory anemia with ring sideroblasts and thrombocytosis (RARS-T) (10). While a few cases of atypical *BCR/ABL*-negative CMML and CML present with *JAK2* mutation, the proliferative aspects of these diseases are due to aberrancies in intracellular signaling pathways such as *RAS/MAPK*. During the revision process, the greatest controversy regarding MDS/MPN was related to the provisional entity RARS-T (10), also known as essential thrombocythemia with ring sideroblasts. It is known that a good number of cases of RARS-T presents *JAK2* V617F mutation or, much less commonly, *MPL* mutation and the presence of these mutations is correlated with a variation of the platelet count. For this reason, the idea to move RARS-T to the MPN category. In addition, considering that the few cases of RARS-T studied by *in vitro* culture, have shown poor colony formation with lack of endogenous growth, features correlated more with MDS, some authorities suggest to classify RARS-T in the MDS/MPN category. RARS-T diagnostic criteria include not further classifiable MPN/MDS criteria associated with RARS morphological criteria, platelet count over 450,000/mm³, and the presence of anomalous megakaryocytes observed in ET or PMF (11). 60% of RARS-T cases shows a *JAK* mutation or a *MPL* mutation. During the revision process,

in addition, the category atypical CML has been renamed *BCR/ABL*-negative atypical CML, to emphasize that it is not merely a variant of CML, *BCR/ABL* positive (10).

Myelodysplastic syndromes (MDS)

Myelodysplastic syndromes are clonal disorders of hematopoietic precursor cell characterized by cytopenias, dysplasia in one or more myeloid lineages, ineffective hematopoiesis, increased apoptosis and an AML-prone course (11). Parameters used for myelodysplastic syndromes diagnosis of <20% blasts in bone marrow (BM) and peripheral blood (PB) are the same of 2003 WHO classification. A percentage of blasts >20% defines AML, whereas a percentage <20% defines MDS. The revised WHO classification of MDS has been refined to allow for a more precise and prognostically relevant sub-classification of patients with unilineage dysplasia who, in the previous scheme, were considered as “unclassifiable” (10). The category of “refractory cytopenia with unilineage dysplasia” has been increased in order to include MDS patients with a percentage of blasts <1% in the PB and <5% in the BM, but showing a percentage >10% of dysplastic cells in a single myeloid lineage. In this group it is possible to find refractory anemia (RA), refractory neutropenia (RN) and refractory thrombocytopenia (RT). The criteria also permit patients to be included in one of these categories if the dysplasia is unilineage but there is bi-cytopenia in the PB. The category of refractory cytopenia with multilineage dysplasia (RC/MLD) includes, in the 2008 WHO classification, RC/MLD with 15% or more of ring sideroblasts, because there are no differences in prognosis of these two entities. Moreover, to emphasize the prognostic

significance of increased blasts in the PB, the revision process redefines the criteria for refractory anemia with excess blasts (RAEB-1), including patients with 2% to 4% blasts in the blood and less than 5% blasts in the BM that don't present other clinical and laboratory findings normally identified in MDS. In addition, the 4th edition includes a section for pediatric MDS (P-MDS) and a provisional category named childhood refractory cytopenia (RCC). In this last category, childhood cases with less than 2% blasts in the blood and less than 5% in the BM and persistent cytopenias associated with dysplasia in at least 2 cell lineages (10) are present.

Acute myeloid leukemia (AML) and related precursor neoplasms

In the 3rd WHO classification only 4 abnormalities involving *RUNX1*, *CBFB*, *RARA*, *MLL* genes were included in the subgroup "AML with recurring genetic abnormalities ". Later, it has become more widely appreciated that multiple genetic lesions cooperate to establish the leukemic process and influence its morphologic and clinical characteristics (10), so they were incorporated in the 2008 WHO classification. Although for a diagnosis of AML blasts percentage must be >20%, for particular entities, that are AML with t (8;21) (q22;q22), AML with inv(16)(p13.1q22) or t(16;16)(p13.1;q22) and APL with t(15;17)(q22;q12), the genetic abnormality is sufficient for the diagnosis of AML in the appropriate setting regardless the blast percentage in peripheral blood or bone marrow. In the present classification, the so-called variant promyelocytic leukemias (PMLs), presenting *RARA* translocations with other partner genes are identified separately; not all of them have typical APL morphology features

and some show all-*trans* retinoic acid (ATRA) resistance. Another challenge respect to the 3rd edition concerns the subgroup of AML with abnormalities of 11q23; MLL, that has been reviewed to underline the most frequent MLL abnormality: t(9;11)(p22;q23) MLLT3-MLL. 11q23 rearrangements with over 80 partners are still termed MLL abnormalities (11), although the AMLs resulting from these variations are not identical. It is in fact recommended that variant *MLL* translocations should be specified. In addition, given the importance of chromosomal rearrangements such as *DEK-NUP21*, *RPN1-EVI1* and *RBM15-MKL1* in AML, these three new cytogenetic entities are present in the actual classification. Moreover, two provisional entities are present: AML with mutated *NPM1* and AML with mutated *CEBPA*. A particular attention is given to *FLT3* mutations, that are not included in a specific category, but, because of its prognostic importance, the mutational status of *FLT3* should always be ascertained in AML, particularly in cytogenetically normal AML (10).

AML with multilineage dysplasia was initially present in the WHO classification to include AML cases showing MDS-like features. Although the prognostic significance of AML with multilineage dysplasia has been verified in some studies, others showed that morphologic multilineage had no independent prognostic significance when cytogenetic findings were incorporated in the analysis (10). Considering these studies, the name of this category has been changed in “AML with myelodysplasia-related changes”, including cases that show 20% or more blasts in the PB or BM, progress from previously documented MDS or MDS/MPN, present specific myelodysplasia-related chromosome changes and show dysplasia in 50% or more in two myeloid lineage cells.

Regarding therapy-related myeloid neoplasms previously encompassing AML, MDS, and MDS/MPN, it now present as a single AML subgroup (11). Although t-AML, t-MDS and t-MDS/MPN may be considered separately depending on the blast count, they can be defined like a single entity, from a biologic point of view. Considering the 2008 WHO classification criteria, it is not necessary to differentiate these entities according to the therapeutic agent used, like alkylating agents and/or radiation.

Acute myeloid leukemia, not otherwise specified (AML, NOS) category includes all cases that don't meet criteria for any of the other AML subgroups. Clearly, this category comprises always less entities, as more genetic subgroups are recognized.

Three new myeloid neoplasms have been considered in the actual classification: Myeloid sarcoma, the myeloid proliferations related to Down syndrome and blastic plasmacytoid dendritic cell neoplasm. Myeloid sarcoma, previously named granulocytic sarcoma or chloroma, consists of extramedullary blasts proliferation from one or more myeloid lineages, replacing the original tissue normal architecture (11). The myeloid proliferations related to Down syndrome (transient abnormal myelopoiesis and myeloid leukemia) show the same morphologic, immunophenotypic, molecular and clinical characteristics in addition to *GATA1* mutation. The neoplasm known like "blastic NK cell lymphoma" is now known as derived from precursors of a specialized subset of dendritic cells, plasmacytoid dendritic cells, and is named blastic plasmacytoid dendritic cell neoplasm.

Acute leukemias of ambiguous lineage

Acute leukemias of ambiguous lineage show no clear evidence of differentiation along a single lineage (10). Trying to clarify the definition of this group of diseases, the new version of the WHO classification allocates these diseases to a separate chapter, distinct from those of AML and ALL, and changes diagnostic criteria used to define these leukemias. Thus, cases with no lineage-specific markers are termed Acute Undifferentiated Leukemia (AUL), usually expressing CD34, HLA-D and/or CD38 and sometimes TdT, but they do not express specific myeloid or lymphoid antigens (11). In addition, leukemia whose blasts coexpress certain antigens from more than one lineage in the same cell or showing different blast populations from distinct lineages are classified as Mixed Phenotype Acute Leukemia (MPAL). These cases, in particular, can be also classified like B-myeloid or T-myeloid, depending on the presence of different blast populations.

Furthermore, cases of *BCR-ABL1*-positive and *MLL*-positive acute leukemias may meet the criteria for MPAL; only in the presence of *BCR-ABL1*-positive disease, CML in blast phase should be excluded (10).

The introduction of a new category has been made, named “Natural killer cell lymphoblastic leukemia/lymphoma”. Although the neoplasm phenotype of true NK cell precursors is not clear, the diagnosis considers blasts expressing CD56 with immature T-associated antigens, like CD7 and CD2, in the absence of any B or myeloid antigen expression, and with no rearrangement of T cell receptor gene (11).

Precursor lymphoid neoplasm: B lymphoblastic leukemia/lymphoma and T lymphoblastic leukemia/lymphoma

In the new WHO classification, the designation B/T precursor leukemia/lymphoma has been replaced by B/T lymphoblastic leukemia/lymphoma, giving particular attention to the presence of genetic changes too. Although the distinction between lymphoblastic leukemia and lymphoblastic lymphoma is obvious when the patient has a mass composed of B or T lymphoblasts and no blasts in the blood or marrow, it is more arbitrary when there is a mass and limited marrow involvement (10).

The WHO 2008 classification considers the convention used for therapeutic approach to distinguish lymphoma from leukemia and suggests that when the patient presents a mass accompanied by 25% or more nucleolate cells in bone marrow (lymphoblasts), the right diagnosis is acute lymphoblastic leukemia rather than lymphoma. In addition, considering that ALL rarely presents with low BM blast counts, if there are <20% blasts, ALL diagnosis should be postponed until irrefutable evidence confirms the diagnosis (11). However, when there is a patient presenting less of 20% of blasts in bone marrow, without an extramedullary mass, but showing recurrent chromosome changes associated with ALL, the diagnosis has to be lymphoblastic leukemia.

In the case of B lymphoblastic leukemia/lymphoma, numerous reports have demonstrated that recurring genetic abnormalities are associated with sufficiently unique clinical, immunophenotypic, and/or prognostic features so that they can be considered as distinct

entities (10). In this context, for example, ALL cases in children with normal karyotype should be complemented by investigating t(12;21)(p13;q22) by fluorescence *in situ* hybridization (FISH) or ETV6/RUNX1 rearrangement (11) undetectable by routine cytogenetic analysis, as the finding of this abnormality corresponds to a favorable prognosis. In the same way, the assessment for *Philadelphia* chromosome presence and/or BCR/ABL rearrangement is necessary, because it correlates with a poor prognosis.

Particular attention in this new revision is given to the category named B lymphoblastic leukemia/lymphoma with t(5;14)(q31;q32); *IL3-IGH*, associated with differential diagnosis of peripheral blood eosinophilia. In the cases with eosinophilia, the presence of the *IL3-IGH* rearrangement even with a relatively low percentage of blasts, is enough to confirm the diagnosis of ALL. Moreover, when faced with eosinophilia, *FGFR1* rearrangements should also be sought, which, if present, describe the diagnosis of leukemia/lymphoblastic lymphoma associated with *FGFR1* rearrangement (11). If the patient doesn't present these specific abnormalities, the right diagnosis is "B lymphoblastic leukemia/lymphoma, not otherwise specified.

A particular attention is drawn also to the specification that the designation of B-ALL must not be used for Burkitt's lymphoma, representing a neoplasm of mature B cells.

The neoplastic cells of 50% to 70% of patients with T lymphoblastic leukemia/lymphoma demonstrate abnormal karyotypes (10), involving the alpha and delta T cell receptor loci at 14q11.2, the beta locus at 7q35 or the gamma locus at 7p14-15. As the pathogenetic significance of such abnormalities is not clear yet, this leukemia subtype is not sub-classified according to the genetic change (11).

Myeloproliferative neoplasms: Chronic Myeloid Leukemia (CML)

Chronic Myeloid Leukemia

Chronic Myeloid Leukemia (CML) is a myeloproliferative disorder of pluripotent hematopoietic stem cells (13). The disorderly expansion of myeloid progenitor cells appears to result from alterations in their proliferative capacity and a shift in the balance between self-renewal and differentiation toward differentiation, increasing the number of progenitor cells and reducing the pool of stem cells.

CML has an annual incidence of 1-2 cases per 100,000 individuals, and this incidence, that is known to increase with the age, has been the same over the last 50 years. CML accounts for approximately 15% of all leukemias and 7-20% of adult leukemias.

CML is probably the most largely studied human malignancy, and the discovery of the Philadelphia (Ph) chromosome in 1960 as the first consistent chromosomal abnormality correlated with a particular type of leukemia has represented a breakthrough in cancer biology.

CML typically evolves along three clinical phases. There is an initial chronic phase, followed by an accelerated phase that eventually transforms into the blastic phase. Initial symptoms are often the result of anemia or splenomegaly, and include fatigue and left upper abdominal pain or mass (14). Regarding the accelerated phase, it presents with increased constitutional symptoms, progressive splenomegaly, cytogenetic clonal evolution, leukocytosis, and

thrombocytosis or thrombocytopenia. In addition, there is evidence of progressive maturation arrest with an increased number of blasts and basophils. Patients who develop blastic phase show weight loss, fever, night sweats and bone pain, as well as infections and bleeding. It is possible to find extramedullary hematopoiesis, involving the lymph nodes, skin, subcutaneous tissues, bone, and central nervous system (14).

Biology of Chronic Myeloid Leukemia

CML is defined by the presence of the Philadelphia (Ph) chromosome or of the BCR-ABL transcript in the leukemic cells (14). The Ph chromosome is a shortened chromosome 22 derived from a balanced translocation between the long arms of chromosomes 9 and 22, $t(9;22)(q34;q11)$ and is present in 90% of CML patients. In the remaining 10% of Ph-negative patients, BCR-ABL transcripts can be found in approximately 30-50% (14). These particular cases, that result Ph-negative but BCR-ABL positive, show the same clinic-pathologic features and prognosis of Ph-positive CML cases, responding in the same way to therapy. In addition, patients really Ph-negative and BCR-ABL negative, represent a heterogeneous group with an inferior prognosis that is known as Ph-negative CML or atypical CML, proliferative variants of myelodysplastic syndromes, or chronic myelomonocytic leukemia (CMML) (15), (16).

The Ph translocation adds a 3' segment of the *ABL* gene from chromosome 9q34 to the 5' part of the *BCR* gene on chromosome 22q11, giving rise to a hybrid *BCR-ABL* gene that is transcribed into a chimeric *BCR-ABL* messenger RNA (mRNA) (17) (Fig1).

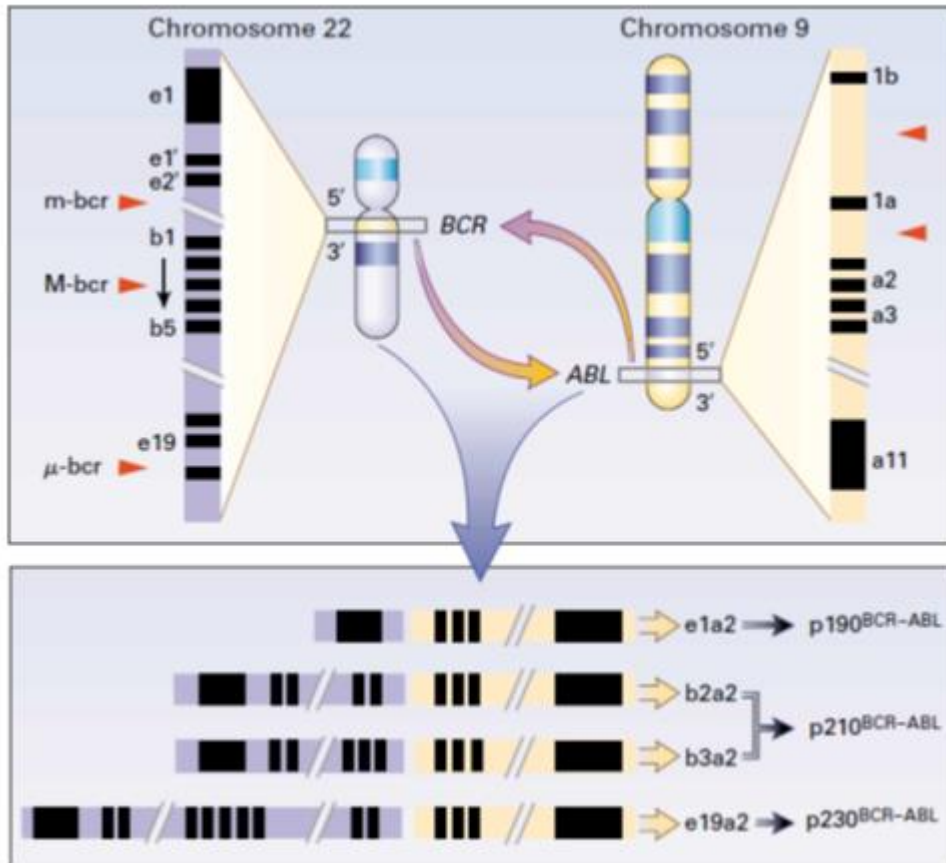


Figure 1. The translocation of $t(9;22)(q34;q11)$ in CML. The Philadelphia (Ph) chromosome is a shortened chromosome 22 resulting from the translocation of 3'ABL segments on chromosome 9 to 5' BCR segments on chromosome 22. Usually breakpoints (arrowheads) on the ABL gene are located 5' of exon a2. Along the BCR gene on chromosome 22 it is possible to identify different breakpoint locations generating different-sized segments of BCR that are subsequently fused with the 3' sequences of the ABL gene. These mechanisms result in fusion messenger RNA molecules (e1a2, b2a2, b3a2, and e19a2) of different lengths that are transcribed into chimeric protein products (p190, p210, and p230), with variable molecular weights and functions. m-bcr: minor breakpoint cluster region; M-bcr: major breakpoint cluster region; μ -bcr: a breakpoint downstream from the M-bcr region between exons e19 and e20. (S. Faderl et al. The biology of Chronic Myeloid Leukemia. N Engl J Med. 1999; 341(3))

The physiologic function of the translocation partners

The *ABL* gene is the human homologue of the *v-abl* oncogene carried by the Abelson murine leukemia virus (A-MuLV); it is normally localized on chromosome 9(q34), has 11 exons and encodes a 140-kilodalton (kDa) protein with weak tyrosine kinase activity. *ABL* gene has two alternate exon 1 sequences that are transcribed differentially from two different promoters (Fig.2). These exons are designated 1b and 1a and their respective promoters are Pb and Pa (14). These two different promoters direct the synthesis of 2 different mRNAs of 6kb and 7 kb, respectively. Human ABL is an ubiquitously expressed 145 kDa protein structurally composed by three SRC-homology domains (SH1-SH3), located toward the NH₂ terminus. The tyrosine kinase function is principally mediated by SH1 domain, whereas the SH2 and SH3 domains are important for the interaction with other proteins. The central part of the protein is characterized by proline-rich regions capable of binding to SH3 domains of other proteins and it harbors 1 of 3 nuclear localization signals (NLS) (Fig.2) (18).

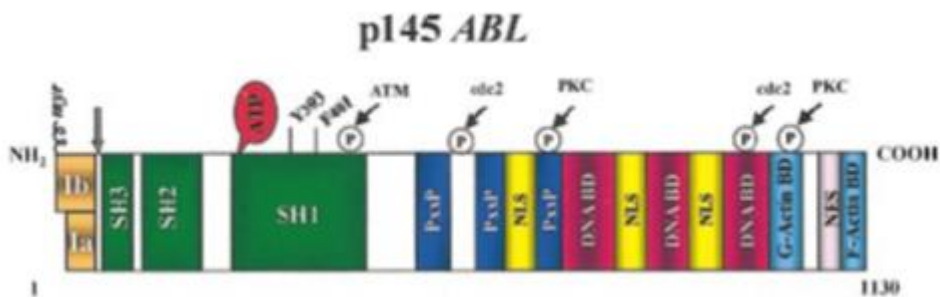


Figure 2. Structure of the ABL protein. The difference between the isoforms 1a and 1b is that the second contains a myristoylation (myr) site for attachment to the plasma membrane. Structurally the

ABL protein is composed from three SRC-homology (SH) domains situated toward the NH₂ terminus; in particular, in the SH1 domain, there is the major site of autophosphorylation within the kinase domain, represented by Y393 and the phenylalanine 401 (F401), a highly conserved site in PTKs containing SH3 domains. In the middle of the protein it is possible to identify different proline-rich regions (PxxP), capable of binding to SH3 domains, and 1 of the 3 nuclear localization signals (NLS). At the carboxy terminus there are DNA as well as G- and F- actin-binding domains. Phosphorylation sites by Atm, cdc2, and PKC are shown. The arrow indicates the position of the breakpoint in the BCR-ABL fusion protein. (M Deininger et al. The molecular biology of chronic myeloid leukemia. Blood 2000;96:3343-3356)

Different functions have been attributed to ABL protein: it is known, in fact, that ABL is involved in the regulation of cell cycle (19), in the cellular response to genotoxic stress (20), and in the transmission of information about the cellular environment through integrin signaling (21).

The *BCR* gene (Breakpoint Cluster Region) is ubiquitously expressed as a 160 kDa cytoplasmatic protein (Figure 3) (18) and is located on chromosome 22(q11).

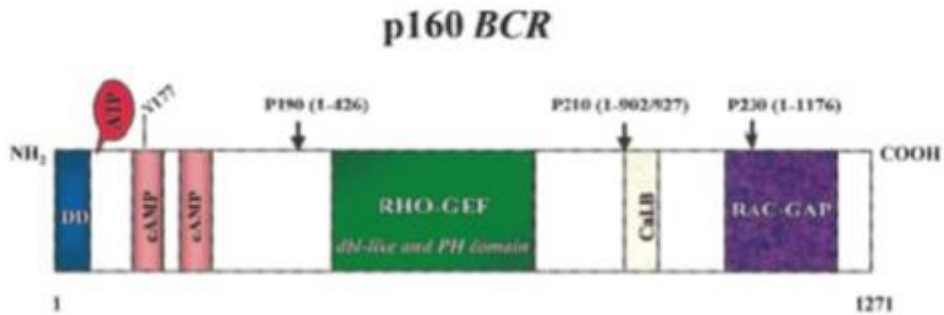


Figure 3. Structure of the BCR protein. Toward the NH₂ terminus are present a dimerization domain (DD) and two cyclic adenosine monophosphate kinase homologous domains. Y177 is the autophosphorylation site crucial for binding to GRB-2. In the middle part there are a region homologous to Rho guanidine nucleotide exchange factors (Rho-GEF) as well as a *dbl*-like and pleckstrin homology (PH) domain. The carboxy terminus contains a putative site for calcium-dependent lipid binding (CaLB) and a domain with activating function For Rac-GTPase. Arrows indicate the position of the breakpoints in the BCR-ABL fusion proteins. (M Deininger et al. *The molecular biology of chronic myeloid leukemia. Blood* 2000;96:3343-3356)

The Bcr region consisting of the first 426 amino acids is encoded by the first exon (e1) of the gene and it is present in all the BCR-ABL isoforms derived from the reciprocal translocation. The N-terminal region of the protein contains a serine-threonine kinase domain and a Tyrosine in position 177 (Tyr177) that, when phosphorylated, forms a binding site for Grb-2, an adapter protein that links BCR to the Ras pathway (22). In the structure of the protein it is possible to identify a dimerization domain (DD) and two cyclic adenosine monophosphate kinase homologous domains at the N-terminal region; in the center, in addition, there is a region homologous to Rho guanidine nucleotide exchange factors (Rho-GEF) as well as *dbl*-

like and pleckstrin-homology (PH) domains that normally are able to stimulate the exchange of guanosine triphosphate (GTP) for guanosine diphosphate (GDP). The C-terminus has GTPase activity for Rac, a small GTPase of the Ras superfamily that regulates actin polymerization and the activity of an NADPH oxidase in phagocytic cells (23). Results obtained from BCR knockout mice suggest that the functions of BCR are not essential for normal cell physiology (24). In fact, BCR-null mice are viable and characterized only by an increased neutrophil respiratory burst.

Molecular anatomy of the BCR-ABL translocation

The reciprocal translocation t(9;22)(q34;q11) between *ABL* and *BCR* leads to the formation of the *BCR-ABL* fusion gene on chromosome 22, generating the Philadelphia chromosome. The mechanisms responsible for this genetic rearrangement remain elusive. The idea is that the localization of chromosome 9 and 22 in the nucleus facilitates the fusion between these two genes during the transition between the S and the G₂ phase of the cell cycle. Using Fluorescent in situ hybridization (FISH) and confocal microscopy, it is possible to identify a juxtaposition of *ABL* and *BCR* during this two phases (25). It has also been suggested that homology regions, such as Alu sequences, between the two chromosomes, may have a role in the *BCR-ABL* recombination that initiates CML (26). From epidemiologic studies it is well known that exposure to ionizing radiation (IR) is a risk factor for CML (27). It is known that IR is able to induce DNA double-strand breaks and it has been shown that high dose of IR can lead to the *in vitro* formation of the *BCR-ABL* fusion transcript in hematopoietic cells.

However, the presence of the *BCR-ABL* translocation in a hematopoietic cell is not in itself sufficient to cause leukemia because *BCR-ABL* fusion transcripts of M-bcr and m-bcr type are detectable at low frequency in the blood of many healthy individuals (28).

Different breakpoint regions on *ABL* and *BCR* genes can create various *BCR-ABL* isoforms. The breakpoints within the *ABL* gene at 9q34 can occur anywhere over a large (greater than 300 kb) area at its 5' end, either upstream of the first alternative exon 1b, downstream of the second alternative exon 1a, or, more frequently, between the two (Fig 4) (18). In all these three cases the fusion protein doesn't change, because the *ABL* exon 1 is never transcribed and the BCR-ABL protein consists of *BCR* exons fused directly to *ABL* exon a2. In contrast to *ABL*, the breakpoints within the *BCR* gene are also found in three different regions but, unlike those in *ABL*, they have a functional impact on BCR-ABL protein, giving rise to three oncoprotein variants. The breakpoint region on the *BCR* gene are known as Minor Breakpoint Cluster Region (m-BCR), Major Breakpoint Cluster Region (M-BCR) and Micro Breakpoint Cluster Region (μ -BCR) (Fig 4) (18) and lead to the formation of the BCR-ABL variants. Since the *ABL* component of the oncoprotein is unchanged, the variability in disease phenotype is largely due to *BCR*.

There are three main BCR-ABL variant proteins (Fig 4):

- p230, the longest form usually expressed at low levels in the leukemic cells and with less oncogenic potential than the other variants, associated with the rare Ph-positive chronic neutrophilic leukemia (29);
- p210, typically associated with CML;
- p190, the smallest form, associated with Acute Lymphoblastic Leukemia (ALL).

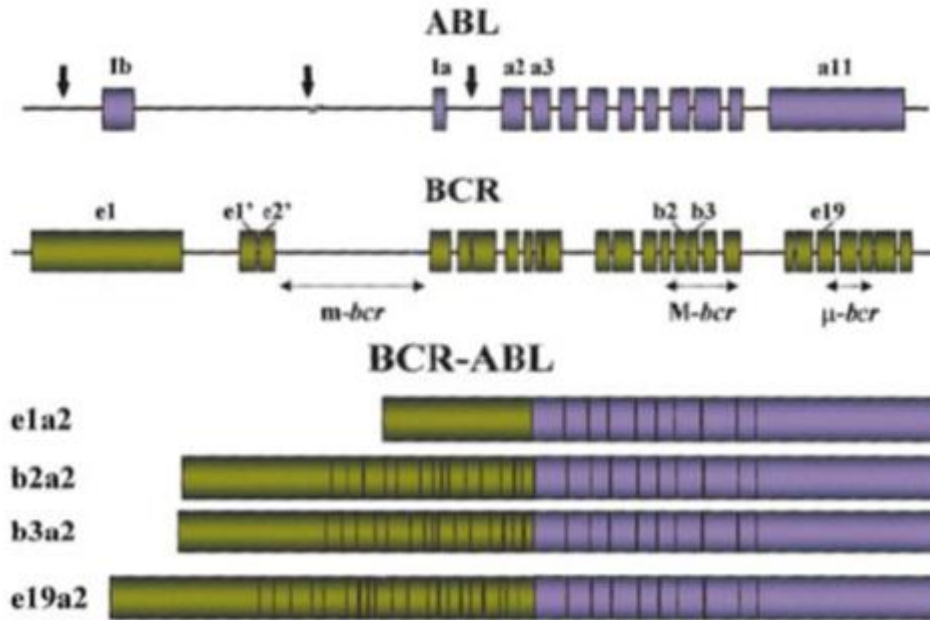


Figure 4. Locations of the breakpoints in the ABL and BCR genes and structure of the chimeric mRNAs derived from the various breaks. (M Deininger et al. *The molecular biology of chronic myeloid leukemia. Blood* 2000;96:3343-3356)

Transfection of bone marrow-derived cell lines with a retroviral vector encoding *BCR-ABL* results in growth factor independency and malignant transformation. The N-terminal segment of *ABL* contains three Src-homology (SH) domains (SH1-SH3 domains), important for the regulation of the tyrosine kinase function of ABL, a catalytic domain, and a myristoylation sequence that connects ABL to proteins of the plasma membrane. Defects in the functional integrity of SH2 decrease phosphotyrosine binding and reduce the transforming capacities of ABL (17) whereas deletion of SH3 facilitates transformation of ABL. In addition, the C-terminus includes a DNA-binding domain, nuclear localization signals and a

binding site for actin. It has been suggested that various structural alterations of ABL and BCR allow the leukemogenic transformation of BCR-ABL. The coiled-coil dimerization motif of BCR, for example, influences its tyrosine kinase activity, increasing it. In addition, BCR interferes with the SH3 domain of ABL and constitutively activates the phosphotyrosine kinase activity of ABL. Important for the regulation of the transforming activity of *BCR-ABL* is also the first exon of *BCR*, that binds to the SRC Homology 2 (SH2) domain of *ABL*. It has been shown, in fact, that this particular binding is necessary for the transforming capacity of BCR-ABL (18).

Activated signaling pathways and biologic properties of BCR-ABL positive cells

Three major mechanisms have been implicated in the malignant transformation by BCR-ABL, namely altered adhesion to stroma cells and extracellular matrix (30), constitutively active mitogenic signaling (31) and reduced apoptosis (32).

Altered adhesion properties

It has been demonstrated the defective adherence of immature hematopoietic CML progenitors to bone marrow stroma elements and extracellular matrix (30). This behavior negatively regulates cell proliferation, facilitating CML cells release into the blood. Normal hematopoietic progenitor cells adhere to the extracellular matrix or to immobilized growth-

regulating cytokines and this attachment happens through cell surface receptors, especially integrins. Recent studies, in fact, report an important role for β -integrin in the interaction between stroma and progenitor cells and that CML cells express an adhesion-inhibitory variant of this particular integrin, that is not present in normal progenitors (33). Considering the normal function of integrins in initiating normal signal transduction and the implication of ABL in the intracellular transduction of such signals, it is clear that the transfer of signals that normally inhibit proliferation is impaired in CML cells, due to the presence of a large pool of BCR-ABL protein in the cytoplasm too. Anyway, despite the evidence that BCR-ABL influences integrin function, it is more difficult to determine the precise nature of the biologic consequences, and, at least, in certain cellular systems, integrin function appears to be enhanced rather than reduced by BCR-ABL (18).

Activation of mitogenic signaling

Ras and the MAP kinase pathways

Ras proteins are small GTPase that act as molecular switches, transducing signals from activated receptors to downstream effectors to regulate cell proliferation, survival and differentiation. RAS small GTPases are activated in many hematopoietic growth factor signaling and in hematological malignancies, but their role in hematopoiesis and leukemogenesis is not completely known. However, several links between BCR-ABL and RAS have been defined. Autophosphorylation of tyrosine 177 provides a docking site for the

adapter molecule GRB-2 (22). GRB2 is a 26-kD protein comprised of a SH2 domain and two SH3 domains; it couples BCR-ABL to SOS, a RAS activator. The need of GRB2-Y177F binding for RAS signaling has been suggested by the evidence that mutation of Y177F inhibits the transforming capacity of BCR-ABL. The importance of RAS-dependent signaling for the phenotype of BCR-ABL expressing cells is supported, in addition, by the observation that the use of antisense strategies, expression of dominant negative molecules or chemical inhibitors of RAS pathway suppresses proliferation and sensitizes cells to apoptosis stimuli. All these data suggest a critical role for RAS in the pathogenesis of CML. There is still dispute as to which mitogen-activated protein (MAP) kinase pathway is downstream of RAS in Ph-positive cells (18). It has been showed that after the stimulation of cytokine receptors such as IL-3 there is an activation of RAS with the subsequent recruitment of the serine-threonine kinase RAF to the cell membrane (34). In this scenario, RAF initiates a signaling cascade through the serine-threonine kinases MEK1/MEK2 and ERK, that gives the activation of gene transcription. The relevance of the MAP-kinase pathway in the BCR-ABL mediated transformation is emphasized by the discovery that the oncoprotein activates also the JNK/SAPK-MAPK pathway (Fig 5) (18) (35). There is also some evidence that p38, the third pillar of the MAP kinase pathway, is also activated in BCR-ABL transformed cells, and there are other pathways with mitogenic potential.

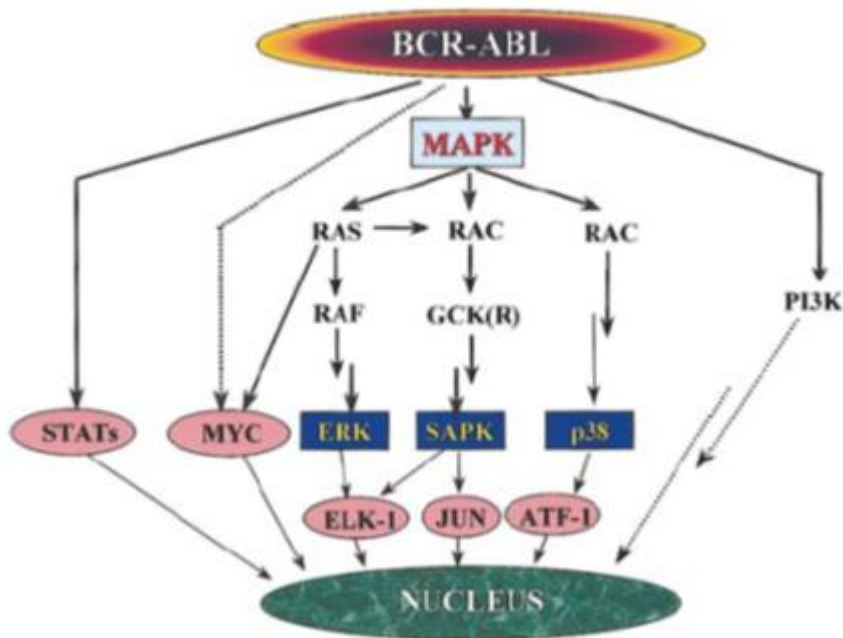


Figure 5. Signaling pathways with mitogenic potential in BCR-ABL-transformed cells. (M Deininger et al. *The molecular biology of chronic myeloid leukemia. Blood* 2000;96:3343-3356)

JAK-STAT pathways

Constitutive phosphorylation of STAT transcription factors (STAT1 and STAT5) has been reported in several BCR-ABL positive cell lines (36) and in primary CML cells (37), and STAT5 activation appears to be necessary for malignant transformation. Although STAT5 acts in different physiological roles, it has been demonstrated that it has an anti-apoptotic effect in BCR-ABL transformed cells, causing transcriptional activation of BCL-xL. In physiologic conditions STAT proteins are phosphorylated downstream of Janus kinase (JAK) activation,

while BCR-ABL appears to directly activate STAT1 and STAT5 without prior phosphorylation. In addition, the importance of the RAS and JAK/STAT pathways in the cellular response to growth factors could explain the observation that *BCR-ABL* renders a number of growth factor-dependent cell lines factor independent (38).

PI-3K/AKT pathway

The PI3k is necessary to permit the proliferation of *BCR-ABL* positive cells (39). It has been demonstrated that BCR-ABL oncoprotein forms multimeric complexes with PI3 kinase and the adapter proteins CBL and CRK, causing the activation of the serine-threonine kinase AKT, particularly important for its anti-apoptotic role. Activation of AKT causes a proliferative effect by down-regulating the cell cycle inhibitor p27 (40) and an anti-apoptotic effect by inducing the phosphorylation of the pro-apoptotic protein BAD, thus preventing BAD-mediated inactivation of the anti-apoptotic protein BCL-xL. Moreover, knowing that usually survival signals mediated by the cytokine IL-3 are transduced into the cell via the PI3/AKT kinase pathway, it is possible that BCR-ABL can mimic the physiologic IL-3 survival signal in a PI3-kinase dependent manner (18).

MYC pathway

Overexpression of MYC has been demonstrated in many human malignancies. It is known to act as a transcription factor and its target genes include genes involved in cell cycle

regulation (e.g. Cyclin-Dependent Kinase 4, CDK4) and genes involved in apoptosis (e.g. prohibitin, also known as PHB).

The link between BCR-ABL and MYC induction is not well understood, although the induction of MYC expression caused by v-Abl has been already described in v-abl transformed cells. The idea is that v-Abl initiates a cascade of phosphorylation involving RAS, RAF, cyclin-dependent kinases (cdks) and the transcription factor E2F that ultimately activates and binds to MYC promoter (41). Similar results were obtained in BCR-ABL transformed murine myeloid cells, but there is no evidence correlating these data with human Ph-positive cells.

Inhibition of apoptosis

Expression of BCR-ABL in factor dependent murine (42) and human (43) cell lines prevents apoptosis after growth-factor withdrawal, an effect that is critically dependent on tyrosine activity and is linked to the RAS activation. In addition, BCR-ABL expressing cell lines have been reported to show an increased resistance to DNA-damage induced apoptosis (32). It has been shown that cells expressing high levels of BCR-ABL show constitutively high levels of p53, p21 and BAX, and high levels of BCL2 and that these levels are not altered by cytotoxic insults. The hypothesis is that when is present at elevated levels BCR-ABL, prevents the translocation of the pro-apoptotic proteins BAD and BAX to the mitochondria inhibiting the release of cytochrome-C and blocking the activation of caspases (44) (45). This effect upstream of caspase activation might be mediated by the BCR-2 family of proteins: in fact, BCR-ABL has been shown to up-regulate BCL-2 in a PI3 kinase-dependent manner in 32D cells

(46). Another link between *BCR-ABL* and the inhibition of apoptosis is represented by the phosphorylation of the pro-apoptotic protein BAD mediated by RAF1; recently it has been demonstrated that the survival signal provided by BCR-ABL is at least in part mediated by RAF1 (47).

Thus, it is clear that multiple proliferative and anti-apoptotic signals initiated by BCR-ABL, are difficult to separate and this is in line with the idea that a proliferative signal leads to apoptosis unless it is counterbalanced by an anti-apoptotic signal (48) and BCR-ABL represents a good example of this theory.

BCR-ABL and the DNA damage and repair

Generally, DNA damage mechanisms can generate single-nucleotide alterations, single-strand or double-strand breaks (DBSs). Usually, double-strand breaks arise from single-strand breaks. Single-nucleotide alterations are repaired by mismatch repair (MMR) or by nucleotide excision repair (NER) mechanisms. Strand-breaks (single and double) are repaired by either high-fidelity homologous recombination when a sister chromatin is available as a template (during the S or G2 phase of the cell cycle), or by non-homologous end joining (NHEJ), giving short deletions in the repaired strands.

It is known that DNA mutations can occur as the result of different situations; it has been reported the presence of mutations in single-nucleotide repair pathways, or in proteins necessary for the process of DNA DSB recognition and repair or, in addition, they can be the product of a failure of cell cycle checkpoints, that causes subsequent replication of damaged

DNA. This last case, in particular, is due to defects in sensing DNA damage or defects in proteins necessary to execute the cell cycle arrest.

The appearance of chromosomal abnormalities in patients with BC-CML has led to many attempts to elucidate the mechanisms by which BCR-ABL affects DNA damage and repair.

Several common themes emerge from previous study (46): it has been demonstrated that the presence of BCR-ABL increases the genomic and chromosomal instability in cells, although the increase appears modest, and after a long period of BCR-ABL expression, the frequency of appearance of DNA damage increments. BCR-ABL has been shown to induce the production of reactive oxygen species, causing oxidative damage and mutation (49).

Effects of BCR-ABL on all the pathways reported have been described and summarized in the figure 6 (50).

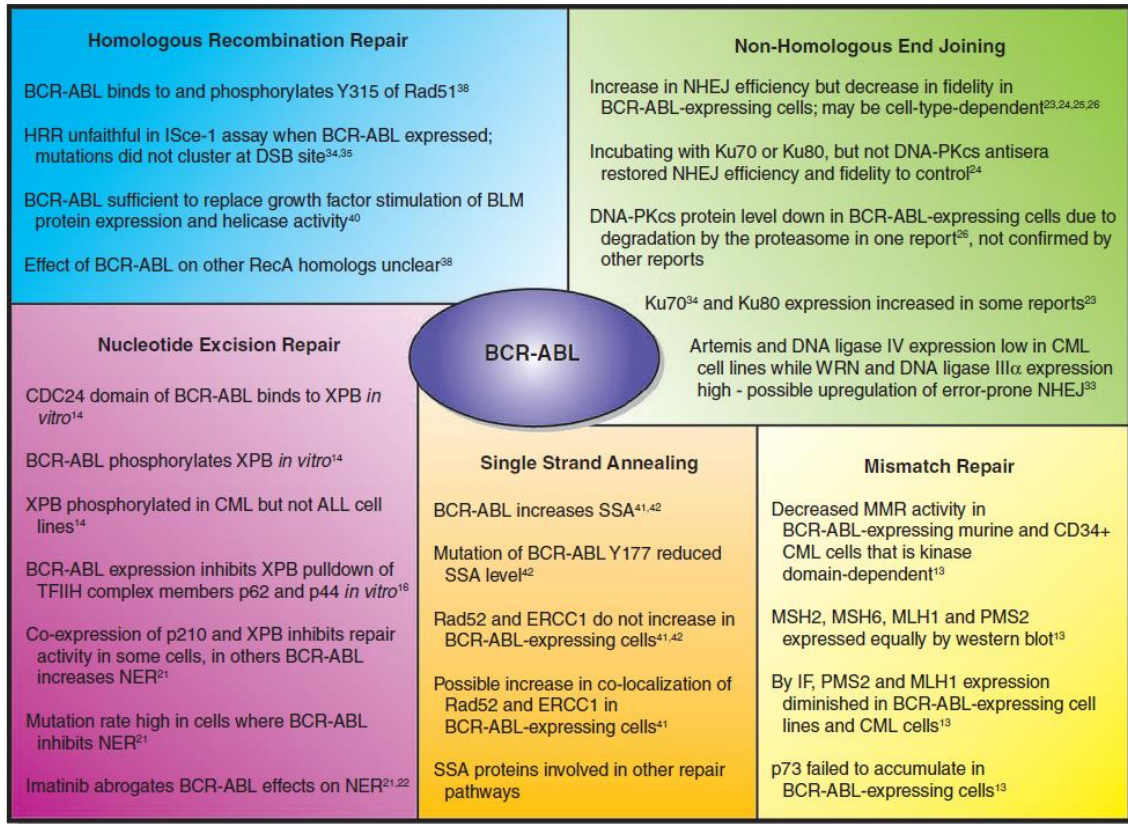


Figure 6. BCR-ABL affects DNA repair processes. (B A Burke & M Carroll. BCR-ABL: a multi-faceted promoter of DNA mutation in chronic myelogenous leukemia. *Leukemia* 2010;6:1105-12)

CML treatment

In CML it is possible to distinguish three different levels of response to the therapy:

- Hematologic response: a Complete Hematologic Response (CHR) means that the number of white cells, red cells and platelets remains unvaried, normal or near normal; Partial Hematologic Response (PHR) means that the number of blood cells is altered;

- **Cytogenetic Response:** Complete Cytogenetic Response (CCyR) means that Ph positive cells are not detectable. The complete response is determined after a cytogenetic analysis, based on a standard karyotype with 20 metaphases that shows the absence of *BCR-ABL* positive cells. Moreover, there are the Major Cytogenetic Remission (MCyR) when the same analysis shows the presence of less than 33% of *BCR-ABL* positive cells and the Minor Cytogenetic Remission (MyCyR), when the percentage of *BCR-ABL* positive cells is more than 33 % and less than 66%. There is absence of cytogenetic remission when more than 90% of *BCR-ABL* positive cells are present (Table 2)
- **Molecular Response:** when there is a reduction of the *BCR-ABL* positive cells there is a partial molecular response; a Major Molecular Response (MMR) means that there is a 1,000-fold decrease in the level of cells with the *BCR-ABL* gene from the level measured at the beginning of treatment. Alternatively, a Complete Molecular Response (CMR) means that the *BCR-ABL* cancer gene is not detectable by PCR (Table 3)

CYTOGENETIC RESPONSE	
Complete Cytogenetic Response (CCyR)	Ph+ cells not detectable
Major Cytogenetic Remission (MCyR)	BCR-ABL + cells <33%
Mynor Cytogenetic Remission (MyCyR)	33%<BCR-ABL+ cells<66%

Table 2. Cytogenetic response in CML.

MOLECULAR RESPONSE	
Complete Molecular Response (CMR)	BCR-ABL cancer gene not detectable
Major Molecular Response (MMR)	1,000 fold ↓ BCR-ABL+ cells

Table 3. Molecular response in CML.

If not effectively treated, CML follows an inexorable course that results in the progression to the advanced phases and eventually in the death of the patient.

The presence of the Philadelphia Chromosome in the Chronic Myeloid Leukemia cells suggested that suppressing the Ph positive clones might change the natural course of this disease and its prognosis.

Hydroxyurea, busulfan and interferon- α were the three agents most commonly used as the standard treatment in patients with chronic myeloid leukemia before the introduction of imatinib treatment, but neither of these drugs produced a substantial, durable suppression of the Ph-positive cells in patients with CML (51), giving the evolution of disease in Blast Crisis phase.

Until the introduction of stem cell transplantation, CML treatments were purely palliative. Currently, only allogeneic stem cell transplantation (SCT) has been shown to provide a long term eradication of the disease (52) but only a small minority of CML patients may benefit of this therapeutic option, because of limitations related to the availability of suitable donors and the treatment related mortality due to acute/chronic graft-versus-host disease or infections (52). Allogeneic SCT is curative in selected patients with CML, and is most effective when performed during the chronic phase (CP) of disease. In CP-CML, in fact, it is associated

with 3-5 year survival rates of 40-80%, and 10 year survival rates of 30-60%. Transplantation-related mortality (TRM) ranges from 5-50% depending on patient age, donor origin, and degree of matching, among other factors (14).

The most exciting of the therapeutic approaches was brought by the advent of signal transduction inhibitors (STI), which block or prevent a protein from exerting its role in the oncogenic pathway (53). Considering that the principal transforming property of the BCR-ABL protein is due to its constitutive tyrosine kinase activity, direct inhibition of this activity appears to be the right way in order to silence the protein. For this aim, different tyrosine kinase inhibitors have been evaluated for their ability to modify the CML cells phenotype. Among these, there were compounds isolated from natural sources or synthetic compounds developed through a rational design of chemical structures capable of competing with the adenosine triphosphate (ATP) or the protein substrate for the binding site in the catalytic center of the kinase (Fig7) (18). The most promising of these compounds was a small molecule known as Imatinib (CGP57148B, STI571; Glivec Novartis Pharmaceuticals), which specifically inhibits ABL tyrosine kinase at micromolar concentrations and has revolutionized the treatment and prognosis of CML. Like many other receptor tyrosine kinases, BCR-ABL is at an equilibrium between two states, an active state and an-auto regulated inactive state. Imatinib is a small molecule 2-phenylaminopyrimidine that acts as an ATP mimic thus occupying the binding site for ATP within BCR-ABL, stabilizing, in addition, the inactive conformation of BCR-ABL, in which the well-known "DFG triad" (the highly conserved aspartate-phenylalanine-glycine motif) is in the "out" conformation and finally causing the inhibition of the phosphorylation of tyrosine residues on substrate proteins and BCR-ABL

itself (54). Inhibition of the BCR-ABL kinase activity by this compound results in the transcriptional modulation of different genes important for the control of the cell cycle, cell adhesion, and cytoskeleton organization and crucial for CML leukemogenesis, inducing the Ph-positive cell to an apoptotic death (55).

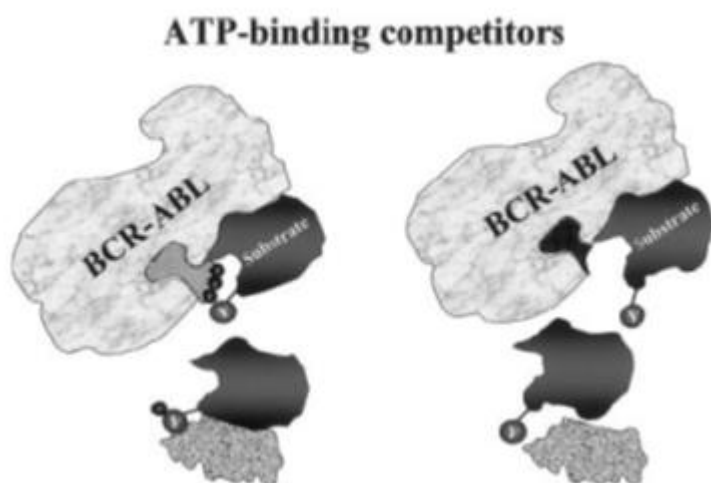


Figure 7. Mechanism of action of tyrosine kinase inhibitors. The drug competes with ATP for its specific binding site in the kinase domain. Whereas the physiologic binding of ATP to its pocket allows BCR-ABL to phosphorylate selected tyrosine residues on its substrates (left), a synthetic ATP mimic as imatinib binds equally this pocket, but not providing the essential phosphate group to transfer to the substrate (right). (M Deininger et al. The molecular biology of chronic myeloid leukemia. Blood 2000;96:3343-3356)

Different studies were focused on the effects of this particular compound on BCR-ABL positive cell lines; these studies reported that Imatinib inhibited the kinase activity of ABL

both *in vitro* and *in vivo*, the growth of v-abl and bcr-abl transfectants, as well as the *in vitro* formation of bone marrow (BM)-derived colonies in the presence of growth factors in some CML patients (55). In addition, cellular proliferation and tumor formation by BCR-ABL expressing cells were specifically inhibited by this compound, that caused a 92-98% decrease in the number of bcr-abl colonies formed in colony-forming assays of PB or BM from patients with CML (54, 56).

Imatinib can also cause regression and eradication of Ph-positive tumors in nude mice (57).

In addition to p210^{BCR-ABL}, imatinib inhibits several other tyrosine kinases including p190^{BCR-ABL}, v-Abl, c-ABL, c-KIT, and platelet-derived growth factor-receptor (PDGF-R) (14).

Imatinib is now established as the front-line therapy for CML (58, 59), despite the fact it does not completely eliminate *BCR-ABL* expressing cells (60).

Some patients treated with imatinib can develop resistance to this inhibitor, especially if treated in advanced phase, and the emergence of resistance, particularly in patients with acute leukemias, has prompted intense research to overcome this problem. Hematological resistance to imatinib can be defined as the lack of a complete hematological response in patients with chronic phase disease (58), a return to chronic phase for patients in acute phase (61), or a partial response in patients with blast-crisis CML or Ph-positive ALL (62). In addition, resistance can be primary (if there is no response to imatinib after the initial treatment) or secondary (when there is resistance after achieving an objective response) (63). Usually patients with acute leukemia (blast-crisis CML or Ph-positive ALL) develop resistance to imatinib, that happens within 3-6 months of treatment in over 70% of them (62). Once resistance develops, the disease is rapid and aggressive in most patients. Two

main mechanisms are supposed to give resistance to imatinib: amplification of *BCR-ABL* gene and mutations in the catalytic domain of the protein (63). It has been shown that resistance develops when cells find ways to maintain sufficient amounts of *BCR-ABL* signaling, and this can be due to the amplification of the *BCR-ABL* gene, to the presence of *BCR-ABL* gene mutations or to an incomplete *BCR-ABL* inhibition, which can favor the selection of resistant cells. The amplification of the *BCR-ABL* gene in imatinib-resistant leukemic cells has been initially described in a LAMA84R cell line generated by culturing cells in increasing concentrations of imatinib (63), and was confirmed in other cell lines, although some differences were noted among them. The *BCR-ABL* gene amplification was observed in clinical samples obtained from patients resistant to imatinib, too, with a frequency of about 18% (63).

Resistance most often originates as a consequence of mutations in the kinase domain of *BCR-ABL*, (63) with one survey noting a 90% correlation (64). These mutations are thought to induce resistance by disrupting amino acids that contact imatinib or by prevention of the formation of the inactive conformational state (64). Several groups have characterized mutations in leukemic cells from patients who had developed resistance to imatinib (65). The two most frequently occurring mutations are those affecting amino acids E255 and T315: in the first case, glutamine is substituted by either a lysine (E255K) or a valine (E255V), in the second threonine is substituted by an isoleucine residue (T315I) (63). The T315I mutation is responsible for 20% of all cases of imatinib resistance (66); this gatekeeper threonine is common to many tyrosine kinases and its mutation has been shown to enhance their kinase and transformation activities (66).

Other minor mechanisms of resistance to imatinib can be due to the overexpression of other tyrosine kinases as the SRC-related LYN kinase or variability in the amount and function of the drug influx protein OCT-1 (67).

Although imatinib represents an established first-line therapy for CML that can produce clinical remissions in most patients (particularly in chronic phase), resistance to this drug may occur, so there is necessity to test new drugs for CML patients (68).

New TKIs are available or in development that are more potent than imatinib and have a decreased potential for resistance. These include dasatinib (Sprycel; BMS-354825, Bristol-Myers-Squibb), an orally bioavailable dual BCR-ABL and SRC inhibitor, and nilotinib (Tasigna; AMN-107, Novartis), a potent selective BCR-ABL inhibitor (69).

Dasatinib is at the forefront of the second generation of anti-leukemia drugs, having broader anti-kinase activity than imatinib, dually inhibiting both BCR-ABL and downstream SRC family kinases. It inhibits proliferation of BCR-ABL positive bone marrow progenitor cells from patients with imatinib sensitive and imatinib resistant disease and is active against the majority of imatinib-resistant mutants, except for T315I mutant (68). It has been shown that imatinib binds only to the inactive conformation of the ABL kinase, while dasatinib is able to interact with ABL in multiple conformation state (active and inactive).

Nilotinib is a close analogue of imatinib with approximately 20-fold higher potency for BCR-ABL kinase inhibition; similarly to imatinib, it binds the inactive conformation of BCR-ABL kinase, but the inclusion of alternative binding groups to the N-methyl-piperazine moiety improves the interaction between the drug and the kinase binding site (69).

There are third-generation TKIs too, as Bosutinib (SKI-606) or Ponatinib (AP24534).

Bosutinib is a third generation TKI, dual SRC and ABL kinase inhibitor. Unlike imatinib and dasatinib, bosutinib does not significantly inhibit PDGFR or c-kit but effectively overcomes the majority of imatinib-resistance-conferring *BCR-ABL* mutations except V299L and T315I (70). Bosutinib is effective in upfront treatment of CP-CML and after failure of first- and second-generation TKIs (71).

For patients who do not respond to first-or second-generation TKIs, allogeneic stem cell transplantation and interferon therapy the treatment with bosutinib is the only currently approved with the potential for life prolongation (72).

Novel compounds such as the aurora kinase inhibitors-including danusertib (PHA-739358), homoharringtonine, and ponatinib successfully target the T315I mutation too; clinical trials are ongoing, in fact, and there are promising results (72). In particular, ponatinib (AP24534) is a potent oral tyrosine kinase inhibitor that blocks native and mutated BCR-ABL, including the gatekeeper mutant T315I, which is equally resistant to other tyrosine kinase inhibitors (73). Ponatinib is the product of a computational and structure-based approach to the design of a small-molecule tyrosine kinase inhibitor, in order to elude the steric obstacle due to the presence of isoleucine residue at position 315 in the T315 mutants (73). Ponatinib has showed significant clinical activity in patients with Ph-positive leukemias that were resistant to the other available tyrosine kinase inhibitors. In particular, it has activity in all of the following situations: against a spectrum of mutants in patients in whom previous therapy had failed, against the T315I mutant and against disease that have been refractory to therapy with multiple tyrosine kinase inhibitors in the absence of detectable *BCR-ABL* mutations (74).

However in advanced leukemia even ponatinib fails to secure durable responses in most patients, due to multiple mutations and/or additional mechanisms.

Atypical Chronic Myeloid Leukemia (aCML)

Guidelines for distinguishing chronic myelomonocytic leukemia and atypical chronic myeloid leukemia

In the WHO classification atypical chronic myeloid leukemia has been considered as a new distinct clinical entity included in the category of *myelodysplastic/myeloproliferative diseases*. This category includes three major disorders: chronic myelomonocytic leukemia (CMML), atypical chronic myeloid leukemia (aCML), and juvenile myelomonocytic leukemia (Table 4)(8).

MDS/MPN DISEASES CLASSIFICATION
Chronic myelomonocytic leukemia
Atypical chronic myeloid leukemia
Juvenile myelomonocytic leukemia
Myelodysplastic/myeloproliferative diseases, unclassifiable

Table 4. WHO classification of the myelodysplastic/myeloproliferative diseases. (JW Vardiman et al. The World Health Organization (WHO) classification of the myeloid neoplasms. Blood 2002; 100:2292-2302)

The criteria for the recognition of these disorders are detailed in the WHO monograph, but some aspects of CMML and aCML are problematic and difficult to distinguish (8). In particular aCML is defined as “less than 10% peripheral monocytes” while CMML diagnosis requires “more than 1000 monocytes/cmm”; therefore all samples with <10% monocytes but >1000 monocytes/cmm are of difficult attribution. The differences between these two entities are based primarily on quantitative and qualitative differences in the presenting differential count and blood film morphology (1). In particular, the discriminating features are represented by the relative counts of basophils and monocytes, the absence or presence of a peak of immature granulocytes, the presence or absence of granulocytic dysplasia and the absolute count of neutrophils and late-maturing granulocytes (1). Considering that these characteristics are not so defined, overlap between these two categories is often inevitable, and the choice depends on the arbitrary cut-off points chosen. In particular, the WHO classification divides CMML into two prognostic categories, CMML-1 and CMML-2, based on the number of blasts in the blood and bone marrow (Table 5)(8).

DIAGNOSTIC CRITERIA FOR CHRONIC MYELOMONOCYTIC LEUKEMIA (CMML)
Persistent peripheral blood monocytosis greater than $1 \times 10^9 /L$
No Philadelphia chromosome or <i>BCR-ABL</i> fusion gene
Fewer than 20% blasts in the blood or bone marrow
Dysplasia in one or more myeloid lineages. If myelodysplasia is absent or minimal, the diagnosis of CMML may still be made if the other requirements are present and:
<ul style="list-style-type: none"> - an acquired , clonal cytogenetic abnormality is present in the marrow cells - the monocytosis has been persistent for at least 3 months and all other causes of monocytosis have been excluded
Diagnose CMML-1 when blasts fewer than 5% in blood and fewer than 10% in bone marrow
Diagnose CMML-2 when blasts are 5%to 19% in blood, or 10% to 19% in marrow, or if Auer rods are present and blasts are fewer than 20% in blood or marrow
Diagnose CMML-1 or CMML-2 with eosinophilia when the criteria above are present and when the eosinophil count in the peripheral blood is greather than $1.5 \times 10^9 /L$

Table 5. Diagnostic criteria for chronic myelomonocytic leukemia (CMML). (J W Vardima et al. The World Health Organization (WHO) classification of the myeloid neoplasms. Blood 2002; 100:2292-2302)

Atypical Chronic Myeloid Leukemia (aCML)

As the name of Atypical chronic myeloid leukemia suggests, this particular disorder is an atypical variant of Chronic Myeloid Leukemia (CML) that lacks the Philadelphia (Ph) chromosome and *BCR-ABL* fusion gene that are considered the hallmarks of classical CML (8).

Little is known about this uncommon disease, whose incidence is about one-two cases every 100 cases of Ph-positive CML (75), but is reported that it is clinically a very aggressive disease.

The literature describing aCML prior to 1987 is difficult to evaluate because many cases were not studied carefully for masked or variant translocations and for *BCR* gene rearrangements. From a clinical point of view, aCML shows both myeloproliferative as well as myelodysplastic features and is characterized by leukocytosis and marked myeloid dysplasia, not observed during the chronic phase of CML. Some cases have an absolute monocytosis but can be distinguished from chronic myelomonocytic leukemia (CMML) by the presence of a higher percentage (>15%) of circulating immature granulocytes (76). Moreover, aCML as defined by the WHO criteria is a *JAK2* V617F negative hematopoietic disorder (2), although the identification of an activating 1849G>T mutation of *JAK2* tyrosine kinase has been recently described in 3 of 16 (19%) patients with aCML (77).

In particular, CMML and aCML are defined as distinct hematological entities but emerging data reveal considerable overlap at the molecular level (78). The most common known abnormality found in CMML and aCML is represented by *NRAS* and *KRAS* mutations, present in approximately one third of cases. Despite the identification of these mutations is not so recent, their significance and correlation with pathogenesis and prognosis remain unclear. More recently, other mutations occurring in tumor suppressor genes such as *TET2*, *RUNX1*, *ASXL1* and *CBL* have been identified in a fraction of CMML and aCML patients, but none of these alterations has been characterized in the particular context of these two entities, giving

no additional information for defining the pathophysiology of these two myelodysplastic/myeloproliferative neoplasms.

The overlap between CMML and aCML at molecular level, together with the well-known problem of an accurate morphological classification and to the absence of specific characteristic markers, makes the discrimination and diagnosis of these two very similar diseases very challenging.

Specifically, concerning aCML, because no specific recurrent genomic or karyotypic abnormalities have been identified, the molecular pathogenesis of this disease has remained elusive and the outcome dismal (median survival of 37 months after diagnosis) (5), with no improvement over the last 20 years. In this context, the identification of specific oncogenic lesions is mandatory, in order to have a better classification and, if possible, in order to identify specific treatments for this disease.

Aim

In the WHO classification atypical chronic myeloid leukemia (aCML) has been considered as a new distinct clinical entity included in the category of mixed myeloproliferative/myelodysplastic disorders, with dysplastic as well as proliferative features at the time of initial diagnosis. It is characterized principally by involvement of the neutrophil lineage with leukocytosis resulting from an increase of morphologically dysplastic neutrophils and their precursors. However, multi-lineage dysplasia is common and reflects the stem cell origin of aCML.

The WHO criteria for the diagnosis of atypical CML are:

- Persistent leukocytosis;
- Lack of the BCR-ABL fusion gene;
- Evidence of marked multilineage dysplasia;
- Monocytosis <10%
- Basophils <2%;
- Immature circulating precursors <10%;
- Bone marrow blast count <20%.

Unlike CML, characterized by the *BCR-ABL* fusion, for which the prognosis was markedly improved by the development of imatinib as a specific inhibitor of the BCR-ABL protein, the molecular pathogenesis of aCML remains unclear and the outcome dismal (median survival of 37 months after diagnosis) with no improvement over the last 20 years.

A few cases of aCML have been reported to demonstrate *JAK2* mutations, but the proliferative aspects of most cases are related to aberrancies in the RAS/MAPK signaling pathways. In addition, approximately 30% to 40% of cases of aCML also exhibit *NRAS* or *KRAS*

mutations. The most common karyotypic changes reported in atypical CML include abnormalities of chromosomes 8, 13, 14, 17, 19 and 21.

Although a number of molecular lesions have been identified in aCML patients, to date the prognostic impact or therapeutic implications of these mutations has not been fully clarified.

High-throughput sequencing has proven to be a powerful tool to identify recurrent, specific genetic abnormalities in solid cancers and leukemias with the aim to better define the molecular pathogenesis of these diseases and to potentially refine their diagnosis and treatments.

Considering aCML as an heterogeneous disease, with a poor outcome, and without a specific cytogenetic marker, the first objective of this study is the identification of new recurrent driver mutations, applying high-throughput sequencing strategy to aCML samples, in order to improve our knowledge of the molecular pathogenesis of this disease.

Material and methods

Patients. aCML patients were identified according to the criteria proposed in the 2008 World Health Organization classification (WHO 2008). Exome sequencing technology was applied to eight individuals with aCML enrolled between 2008 and 2011, presenting white blood cell counts included in a range between 22.4 and 89×10^9 cells/l, with ages comprised between 49 and 83 years. Seven patients were male and five were smokers or former smokers.

Cells and cell lines. Bone marrow (BM) and peripheral blood (PB) samples were collected at diagnosis in patients affected by aCML or other hematological disorders. In seven cases bone marrow samples were used; in one case (patient 4) peripheral blood cells. Leukemic cells were isolated using separation on a Ficoll-Paque Plus gradient (GE Healthcare). Myeloid cells were assessed using surface markers (CD33, CD13, or CD117) with fluorescence-activated cell sorting (FACS) analysis. Only samples with a population of myeloid cells >80% were further evaluated. Lymphocytes were obtained by culturing cells initially with 2.5 μ g/ml Phytohemagglutinin-M (PHA-M, Roche) and 200 International Units/ml interleukin-2 (IL-2, Aldesleukin, Novartis) (3-4 days), and subsequently with IL-2 alone (2-3 weeks) and were used as control cells. Only samples with at least 80% lymphoid cells, as assessed by FACS analysis using lymphoid cells markers (CD3, CD4, CD5, CD8 or CD19), were used for subsequent analyses.

Two cell lines were used to perform the in vitro experiments. In particular, the TF1 human erythroleukemia cell line was maintained in RPMI 1640 medium, supplemented with 10% FBS, 2mM L-glutamine, 100 U/ml penicillin G, 80 μ g/ml gentamicin, 20 mM HEPES and 2ng/ml human GM-CSF. The second one was represented by 293T human embryonic kidney cell line; these cells were used for immunofluorescence experiments and were grown in

DMEM medium adding 10% FBS, 2mM L-glutamine, 100 U/ml penicillin G, 80 µg/ml gentamicin and 20 mM HEPES.

Exome sequencing. 1 µg of genomic DNA (extracted using the Invitrogen PureLink Genomic DNA kit) was used to generate all exome libraries. After the fragmentation of genomic DNA to obtain a size of 500 bp, these fragments were processed according to the standard protocol for the Illumina TruSeq DNA Sample Preparation kit, selecting fragments of 200-300 bp in size using 2% agarose gel. Multiplexed genomics libraries were then enriched with the Illumina TruSeq Exome Enrichment kit. Libraries were subsequently sequenced on an Illumina Genome Analyzer Iix with 76-bp paired-end reads using Illumina TruSeq SBS kit v5.

Bioinformatics. Image processing and base calling were performed using Illumina Real Time Analysis Software RTA v1.9.35. Qseq files were deindexed and converted to the Sanger FastQ file format using in-house scripts. Using the Burrows-Wheeler-based BWA alignment tool within the Galaxy framework, the alignment between FastQ sequences and the human genome database (NCBI Build 36/hg18) was performed. The reads matching the reference human genome were more than 90%, with mean exon coverage of >70-fold and the percentage of exons with a mean coverage of $\geq 20x$ over 90% for both the leukemic and control samples. In addition, nucleotides targeting exonic regions and exonic regions plus the surrounding 100 bp were 48% and 68%, respectively, with a resulting overall 28-fold enrichment considering exonic versus non-exonic regions. The alignment files in the SAM format were processed using SAMtools alignment processing utilities (79) after an initial filtration by proper-pair and the conversion in the binary BAM alignment format. Duplicates were removed using the SAMtools rmdup command. After, there was the conversion of BAM

files in to the Pileup format; Pileup data generated from paired cancer and control samples were cross-matched using a dedicated in-house software tool. There were different steps in this process: initially there was the analysis of each data set, after the extraction of the information pertaining to each mismatch, either single nucleotide or indel, together with the corresponding read and mapping quality, read coverage of the mutated locus and specific coverage of each mutation. This intermediate information was stored in a condensed-Pileup format. Subsequently there was the cross-match and filtration of the two condensed data sets considering different parameters as absolute coverage of each position (≥ 20), relative coverage of each variant (≥ 0.35), mapping quality (Phred mapping quality threshold=30) and read quality (Phred read quality threshold=30). The last step was the creation of a dedicated statistical model taking into account the coverage of each variant and the overall coverage in the cancer and control samples as well as the sequence of the reference genome in order to perform variant calling. These variants were usually stored and processed with the aim to predict the effect of nucleotide changes on the normal protein function. Considering the minimum relative coverage of each variant (0.35) and the percentage of leukemic cells in our experiments (>80%) and assuming that the mutations are all heterozygous alterations, the detection of a mutation in 35% of reads means that it is present in 70-87.5% of leukemic cells.

Deep sequencing. 200 ng of cDNA were amplified using the Expand High-Fidelity PCR System (Roche). Each amplicon was gel purified using a 3% agarose gel with the QIAquick Gel Extraction kit and then processed according to the standard protocol for the Illumina TruSeq

DNA Sample Preparation kit. The sequencing of the libraries was performed with an Illumina Genome Analyzer IIx with 76-bp paired-end reads using the Illumina TruSeq SBS kit v5.

Evaluation of copy-number alterations. Mean exonic coverage was calculated for all exons in the Consensus Coding Sequence (CCDS) exonic database in case and control samples. This information was initially used to calculate the median whole-exome coverage in cases and in controls. The mean exonic coverage of each exon was subsequently normalized accordingly. The mean normalized exon coverage was further modified by adding an arbitrary factor (20) to smoothen the effect of very-low-coverage values. Individual case-control \log_2 ratios were then calculated for all the exons in the data set and plotted. The presence of copy-number alterations was detected using a combined approach involving a set of statistical Wilcoxon signed-rank tests performed on sliding exonic windows combined with dedicated heuristic algorithms.

Analysis of recurrent mutations. Sanger sequencing was performed for the following genes: *NRAS* (exons 2 and 3), *KRAS* (exons 2 and 3), *TET2*, *EZH2*, *CBL* (exons 8 and 9), *ASXL1*, *IDH1* (Arg140 codon), *IDH2* (Arg132 codon), *WT1*, *SUZ12*, *RUNX1*, *RBBP4*, *NPM1*, *JARID2* (exons 1-18), *JAK2* (Val617 codon), *EED* (exons 2-12), *DNMT3A* (exon 23) and *CEBPA*. The primers listed in Table 6 were used to sequence *ETNK1* (exon 3), *EPHB3* (exons 3, 6-8, 10 and 11), *GATA2* (exons 5-7), *IRAK4* (exons 8-10), *MTA2* (exons 4-6, 8, 9, 14, 15) and *SF3B1* (exons 14 and 15).

Primer name	Primer sequence (5'->3')
SF3B1_Ex14_F	CCAACATGACTGTCCTTTCTT
SF3B1_Ex14_R	GGGCAACATAGTAAGACCCTGT
SF3B1_Ex14_R_Seq	CAAGATGGCACAGCCCATAA
SF3B1_Ex15_F	TTGGGGCATAGTTAAAACCTG
SF3B1_Ex15_R	AAATCAAAAGGTAATTGGTGGA
EPHB3_Ex3_for	TTCTACTACGAGGCTGACAGCG
EPHB3_Ex3_Rev	ACTGGGACTCCTTGGCAGCT
EPHB3_Ex6_For	TCTGAAGTGCCCACTACGC
EPHB3_Ex8_Rev	CTGAGGCAGACGATAGCGA
EPHB3_Ex10_for	ATGAGGCTGTTCTGGGAGTTTG
EPHB3_Ex11_Rev	GTGAGCCCAATCCAACCTCTCT
ETNK1_Ex3_for	TTAAACAGGCTAATAGCTCGTC
ETNK1_Ex3_rev	CCTATCATAGAGCAAAGCAATGT
GATA2_Ex5_For	CTGGACTCCCTCCCGAGAACTT
GATA2_Ex5_rev	TGTGCCCGCTCCTACCACTCT
GATA2_Ex6_for	GGCACCTGTTGTGCAAATT
GATA2_Ex7_rev	CAGCTCCTCGAAGCACTCC
IRAK4_Ex8_for	TTATTTGTCACAGGTGTCAACA
IRAK4_Ex9_rev	AAATTGATGCCATTAGCTGC
IRAK4_Ex10_For	TCAGTGCAAATATCTTACTGGATG
IRAK4_Ex10_rev	CCACACCAAAGCTGTAAATATCAGA
MTA2_Ex4_For	TTGAAGAGGAATCAAAGCAGC
MTA2_Ex6_Rev	TCTGGGATCTCAGCTTGGTAT
MTA2_Ex8_For	CTAGATTGTAGCAGCTCCATTC
MTA2_Ex9_Rev	TGTCATCTCTCCCACTTCTCC

MTA2_Ex14_for	ACTTTCCTGCTTCAGACCACAA
MTA2_ex15_rev	CTGATGCCAACCTCACTGGT

Table 6. List of primers used for the analysis of recurrent mutations. (R Piazza et al. Recurrent SETBP1 mutations in atypical chronic myeloid leukemia. Nat Genet 2013;45:18-24)

Validation of mutations. Genomic DNA extracted (using the PureLink Genomic DNA Mini kit) from the peripheral blood or bone marrow samples of patients with CMML, aCML, CNL, JMML and unclassified MDS/MPN patients was amplified and sequenced using primers reported in Table 7, designed to cover the complete *SETBP1* coding sequence. The region encoding the SKI homologous domain was sequenced in all samples using primers SETBP1_E_for, SETBP1_E_rev, SETBP1_F_for, and SETBP1_F_rev. The Taq used to perform the PCR amplification was the FastStart Taq DNA polymerase, starting from 100 ng of genomic DNA as template. All the mutations found in aCML samples were validated using PCR amplification and confirmed by traditional Sanger sequencing.

Primer name	Primer sequences (5'-3')
SETBP1_A_for	AACACCTCATCGTGTCTCCATC
SETBP1_A_rev	ACCATAAGACAGGGCTTCCTAC
SETBP1_B_for	AAGTCCATTGCTGGTCAGTTG
SETBP1_B_rev	CTTGTTGACAACCAAACCTCTGG
SETBP1_C_for	GCTTCAACATGCTCATCTTTGTTTCT
SETBP1_C_rev	ACTATCTGCGGAATAACCTTCCCTT
SETBP1_D_for	GGATTGGGTCAAGAATGCCAG
SETBP1_D_rev	TCCCTCATGAATCGTCTCGACTG
SETBP1_E_for	CAACACAGACAGTCTTACTGTGATC
SETBP1_E_rev	TTGGTTGGAAGAGCACTGATG
SETBP1_F_for	CCACTTTCAACACAGTTAGGTG
SETBP1_F_rev	TCTCGTGGTAGAAGGTGTAAGTC
SETBP1_G_for	TTCCAAGTTCAGAATCTCCCA
SETBP1_G_rev	AAGCCTGTGGCTTTGCCTG
SETBP1_H_for	AGCATAAGCACAAGGAAGACCG
SETBP1_H_rev	CAGATGTGTCTGAGGTGCAAAGC
SETBP1_J_for	ACCATTTCTCAGATCATGTCC
SETBP1_J_rev	TTTCTCCACATTCATCTTGCTT
SETBP1_L_for	AGGCTGAGTTGAAGGCACCTTGC
SETBP1_L_rev	CTAGGGAAGGACCTCGCTCTCACT

Table 7. List of SETBP1 primers used for validation and for site-directed mutagenesis. (R Piazza et al.

Recurrent SETBP1 mutations in atypical chronic myeloid leukemia. Nat Genet 2013;45:18-24)

Polymorphism analysis. All the variants identified with exome or Sanger sequencing were filtered-out using the dbSNP135 single nucleotide polymorphisms (SNPs) database. In addition, to assess if *SETBP1* variants were real somatic mutations or previously unreported or rare SNPs, direct sequencing of the corresponding exons of *SETBP1* was performed in a panel of 112 healthy donors. All the variants identified in the healthy donors were considered as SNPs and discarded.

Statistical analysis. To proceed with the statistical analysis two-sided methodologies were used with a significance level of $\alpha=0.05$. Continuous variables were described according to group classification (defined by wild-type and mutated *SETBP1*) by mean, median, 95% CI, mean standard error, and minimum and maximum values and compared across groups by the Wilcoxon test. Categorical variables were described according to groups by the proportion of subjects falling into each category. To compare the proportions was used the χ -square test and to describe continuous variables according to groups were used bar graphs. The study of survival probabilities according to group classification was made using the Kaplan-Meier method and the null hypothesis of equality for the survival function across groups was tested by the log-rank test. The univariate Cox model was used to perform the estimation of the ratio of instantaneous hazards between groups and the multivariable Cox model was used to adjust the possible effect of *SETBP1* mutation for confounders, such as age, sex, white blood cell count, hemoglobin concentration, platelet number and the percentage of peripheral blood blasts.

2-hydroxyglutarate quantification. Cells (2×10^6) were suspended in 80% methanol, centrifuged, dried and stored at -80 °C. The determination of 2-hydroxyglutarate levels was obtained using ion-paired reverse-phase liquid chromatography coupled with negative-mode electrospray triple-quadruple mass spectrometry, and integrated elution peaks were compared with 2-hydroxyglutarate standard curves to obtain absolute quantification (80).

SETBP1 cloning and transfection experiments. A plasmid encoding the long isoform of *SETBP1* cDNA (SC114671, Origene) was used as a substrate for PCR amplification (Expand High Fidelity PCR System, Roche). Clone SC114671 (NM_015559.1) codes for the *SETBP1* variant lacking 54 amino acids at the N terminus compared to the longest *SETBP1* variant (NM_015559.2), so that the Gly870 codon in the longest variant corresponds to the Gly816 codon in the SC114671 clone. To be in line with other reports, coordinates corresponding to the NM_015559.2 isoform were used throughout this work. The amplification of *SETBP1* was obtained using two primers spanning the entire *SETBP1* cDNA and introducing two artificial sites for KpnI and XhoI at the 5' and 3' ends of the coding sequence, respectively. The resulting amplicon was cloned into the p-EntrI Gateway entry vector (Life Technology) and subsequently *SETBP1* molecule was sub-cloned into the pcDNA6.2/N-EmGFP-DEST destination vector using the Gateway clonase system (Life Technology).

293T cell line was transfected with 10 µg of plasmid DNA using Fugene Transfection Reagent (Roche) and the cell selection was performed using 10 µg/ml blasticidin. The isolation of cells expressing wild-type *SETBP1* or *SETBP1* Gly870Ser fused with GFP was obtained after a sorting using a FACSAria (BD Biosciences) flow cytometer.

Mutagenesis. To generate a plasmid encoding for SETBP1 Gly870Ser , specific primers were designed and used to mutagenize the entry vector using the Pfu Ultra High Fidelity enzyme (Agilent); after the digestion of the resulting product with DpnI, 2ul were used to transform the TOP10 competent bacterial strain (Life Technology). Traditional Sanger sequencing was used to confirm the presence of the mutation.

Retroviral infection. MIGR1-EGFP vector was used as a destination vector to clone the sequences encoding wild-type SETBP1 and SETBP1 Gly870Ser after their excision from the pENTR1A. In particular, SETBP1 was excised using the SalI and XhoI restriction enzymes and cloned using XhoI. 10µg of MIGR1-*SETBP1* (encoding wild-type or Gly870Ser protein) or empty MIGR1 vector were used to transfect phoenix packaging cells using FuGENE 6 (Promega) and the resulting retroviruses were collected after 3 days of culture. To obtain TF1 cells stably infected with the SETBP1 retrovirus, 5×10^4 cells were transduced by spin infection in retroviral supernatants supplemented with 4 µg/ml polybrene (Sigma-Aldrich) and 20% RPMI 1640. After 48 hours, the cells expressing wild-type SETBP1 and SETBP1 Gly870Ser were re-suspended in complete medium.

Immunoblotting. TF1 cells (1×10^7), used to perform immunoblotting experiments, were washed with PBS and re-suspended in lysis buffer. The lysis buffer was composed of 0.025 M Tris (pH 8), 0.15M NaCl, 1%NP-40, 0.01 M NaF, 1mM EDTA, 1mM DTT, 1mM sodium orthovanadate and protease inhibitors. Cell lysates were centrifuged for 20 min at 18,000 g. After the quantification, equal amounts of total protein were separated by 10% SDS-PAGE and probed with selected antibodies. The antibodies used in the blotting experiments were: α-SETBP1 (ab98222), α-phosphorylated PP2A (Tyr307, clone E155) (Abcam), α-SET (clone F-9,

Santa Cruz Biotechnology), α -PP2A (C subunit, clone 1D6, Millipore) and α -actin (a2066, Sigma-Aldrich).

Peptide pulldown assay. Biotinylated phosphorylated peptides encompassing the SETBP1 region (amino acids 859-879) of either wild-type or Gly870Ser protein were synthesized by Innovagen (81). To perform the pull-down experiments, 0.5 mg of streptavidin magnetic beads (Pierce Biotechnology) were washed and re-suspended in 100 μ l of TBS supplemented with 2% BSA. The binding between each peptide (10 μ g) and the beads occurred after 1 hour at room temperature on a rotating device; as a negative control, a sample with no peptide was used. According to the protocol, beads were washed with wash buffer, composed of TBS with 0.1% tween 20, and re-suspended in lysis buffer. Subsequently the TF1 cell lysate (600 μ g) was added and incubated for 2 hours at 4 $^{\circ}$ C on a rotating device. Alternatively, 75 ng of recombinant SKP1 CUL1 (SCF) complexed to β -TrCP1 (Millipore) was used. The unbound fraction was used as a loading control and bound proteins were eluted with 30 μ l of Laemmli buffer. The de-phosphorylation of peptides was obtained adding 20 U calf intestinal phosphatase (NEB) after an incubation for 2 hours at 37 $^{\circ}$ C, before binding to streptavidin beads. For the immunoblotting experiments an antibody raised against β -TrCP1 (clone H-85, Santa Cruz Biotechnology) was used.

PP2A activity assay. The PP2A IP Phosphatase Assay kit (Millipore) was used to assess the activity of PP2A using the manufacturer's protocol on 5×10^6 cells.

Proliferation assay. To proceed with the proliferation assay, TF1 cells expressing GFP fused with wild-type or Gly870Ser SETBP1 were seeded at a concentration of 5,000 cells per well in

a 96 well round-bottom cell culture plates using complete medium. After different incubation times (24, 48, 72, and 96 hours) ^3H -methyl-thymidine (Perkin Elmer) was added to the medium and incubated for 8 hours at 37 °C. The cells were then harvested onto glass fiber filters using a Tomtec automated cell harvester and [^3H] thymidine incorporation was measured using a filter scintillation counter (1430 MicroBeta). Each experiment was performed in quadruplicate and was repeated at least twice.

RNA sequencing. All RNA libraries were generated from 2 μg of total RNA extracted with TRIzol (Life Technology) using the standard protocol. RNA was processed according to the protocol for the Illumina TruSeq RNA Sample Preparation kit with a modification in the fragmentation time: mRNA was sheared for 1 min at 94 C°, and, after ligation of the adapters, fragments of 400-500bp were selected on 2% agarose gels. Libraries were sequenced on an Illumina Genome Analyzer Iix with 76-bp paired-end reads using Illumina TruSeq SBS kit v5.

Transcriptome profiling. The software used to perform image processing and base calling was Illumina Real Time Analysis Software RTA v1.9.35. Qseq files were deindexed and converted into the Sanger FastQ file format using in-house scripts (81). As reference genome for the alignment of FastQ sequences was used the human genome database (NCBI Build 36/hg18) and the process was performed using TopHat (82) (version 1.2.0) using default parameters. Reads were mapped using the gene and splice-junction models provided in the Human Ensembl annotation file (Homo_Sapiens.NCBI36.54.GTF) (81). Using an algorithm named Bowtie (83) it was possible to align the RNA-seq reads across the genome and to map the initially unmappable reads to the known splice-junction sequences supplied by the annotation GTF file. In addition, with the Integrated Genomic Viewer (84) we confirmed the

presence of the longer isoform of *SETBP1*, including the SKI homologous region in both wild-type and mutated samples. The quantitative gene expression profile was estimated by SAMMate (85) (version 2.6.1) using Human Ensembl annotation file version 54 and the default parameters. FPKM values, that represent normalized measures of exonic read density and measures of the concentration of a transcript, were used to obtain gene expression values for paired-end data. The FPKM values relative to each gene were calculated by SAMMate, considering the reads mapping on exons or on exon-exon junctions. The Human Ensembl gene annotation file version 54 was used to infer gene expression for coding and non-coding transcripts (81). Using SAMMate it was possible to obtain, having the read alignment information, stored in the BAM format, a matrix of FPKM expression values and read counts for 36,655 unique Ensembl genes. To focus on the gene expression profile of known coding transcripts, a data set of 20,907 protein-coding Ensembl genes was selected from the whole transcriptome (81). Using the DESeq (86) algorithm (R package, version 1.8.1) and NOISeq (R method, version last modified 29 April 2011) (87) was possible to obtain the differential expression profiles between samples with wild-type and mutant *SETBP1*: the FPKM values ranged from 0 to 75,497 and 37% of protein-coding Ensembl genes showed FPKM values between 0 and 1. In order to exclude from the analysis differentially expressed genes biased from very-low-count data or expression values, a filtration was performed, not considering genes having the maximum FPKM value of the two groups <1 , obtaining 13,106 FPKM values. A value of fold change ≥ 3 of the protein coding genes was chosen to focalize the attention on differentially expressed genes, and out of these 1,465 genes, only 197 showed differential expression in samples with wild-type and mutated *SETBP1*. The list of

differentially expressed genes (Table 8) was annotated using the Ensembl gene annotation file (version 54) and IPA (Ingenuity Systems) and the list of TGF- β -related genes (Table 8) was generated by using the Ingenuity Knowledge database. Finally, the FPKM expression values of the differentially expressed gene list were plotted in a heatmap using the dChip software. The 60th percentile value was used as the upper bound of FPKM expression values, and each FPKM was converted into a color scale from -1.5 (lower limit) to 1.5 (upper limit) (81). Data generated from RNA-seq were also used in order to discover possible fusions, using FusionAnalyser software (88).

Gene ID	Gene	Chr	FPKM MeanWT	FPKM MeanM	FPKM Fold Change	pval	padj	P	Label
ENSG00000165029	<i>ABCA1</i>	9	3.29	10.07	3.06	8.66E-03	9.56E-02	0.5	up
ENSG00000214633	<i>AC022400.9</i>	10	0.89	5.91	6.67	4.42E-03	8.77E-02	0.6	up
ENSG00000219041	<i>AC027612.6-2</i>	2	0.17	1.44	8.63	9.95E-03	9.56E-02	0.4	up
ENSG00000196526	<i>AFAP1</i>	4	0.6	2.11	3.5	1.16E-02	9.77E-02	0.3	up
ENSG00000153207	<i>AHCTF1</i>	1	5.55	21.67	3.9	6.60E-03	9.38E-02	0.6	up
ENSG00000187134	<i>AKR1C1*</i>	10	0.54	5.23	9.72	6.48E-03	9.38E-02	0.6	up
ENSG00000186316	<i>AL117190.6-1</i>	14	35.41	2.97	-11.91	5.11E-02	2.64E-01	0.8	down
ENSG00000214515	<i>AL117190.6-3</i>	14	39.24	2.95	-13.31	3.14E-02	1.80E-01	0.8	down
ENSG00000162551	<i>ALPL*</i>	1	1.27	33.6	26.5	6.16E-02	3.03E-01	0.8	up
ENSG00000176597	<i>B3GNT5</i>	3	9.47	76.68	8.09	7.05E-04	8.69E-02	0.7	up
ENSG00000176788	<i>BASP1</i>	5	42.47	173.64	4.09	8.48E-03	9.56E-02	0.6	up
ENSG00000140379	<i>BCL2A1</i>	15	58.88	649.57	11.03	5.42E-04	8.69E-02	0.8	up
ENSG00000134107	<i>BHLHE40*</i>	3	21.92	69.83	3.19	1.04E-02	9.56E-02	0.5	up
ENSG00000125845	<i>BMP2</i>	20	0.82	3.83	4.65	9.05E-03	9.56E-02	0.5	up
ENSG00000159388	<i>BTG2</i>	1	115.45	431.22	3.74	5.60E-03	9.09E-02	0.6	up
ENSG00000166920	<i>C15orf48</i>	15	0.11	11.6	106.31	3.71E-02	2.07E-01	0.8	up
ENSG00000205710	<i>C17orf107</i>	17	0.78	2.53	3.22	1.21E-02	9.86E-02	0.3	up
ENSG00000198723	<i>C19orf45</i>	19	0.03	1.59	53.25	1.09E-03	8.69E-02	0.5	up
ENSG00000178803	<i>C22orf45</i>	22	2.61	15.87	6.08	8.04E-04	8.69E-02	0.6	up
ENSG00000197405	<i>C5AR1</i>	19	54.97	442.32	8.05	3.04E-03	8.77E-02	0.8	up
ENSG00000213085	<i>CCDC19</i>	1	0.96	12.91	13.45	4.20E-04	8.69E-02	0.7	up
ENSG00000205017	<i>CCL3L1*</i>	17	3.12	100.59	32.23	1.70E-03	8.77E-02	0.9	up
ENSG00000205021	<i>CCL3L3*</i>	17	2.37	66.37	27.96	3.12E-03	8.77E-02	0.9	up
ENSG00000205020	<i>CCL4L1</i>	17	4.24	55.76	13.15	1.08E-02	9.56E-02	0.8	up
ENSG00000197262	<i>CCL4L2</i>	17	4.99	75.81	15.2	1.67E-02	1.17E-01	0.8	up

Gene ID	Gene	Chr	FPKM MeanWT	FPKM MeanM	FPKM Fold Change	pval	padj	P	Label
ENSG00000138764	<i>CCNG2</i>	4	17.49	86.69	4.96	3.77E-03	8.77E-02	0.7	up
ENSG00000183625	<i>CCR3</i>	3	0.59	6.66	11.31	2.41E-03	8.77E-02	0.6	up
ENSG00000177575	<i>CD163*</i>	12	115.85	10.62	-10.91	3.05E-01	9.82E-01	0.8	down
ENSG00000204936	<i>CD177</i>	19	20.47	85.49	4.18	1.07E-02	9.56E-02	0.6	up
ENSG00000112149	<i>CD83</i>	6	14.71	144.35	9.81	3.82E-03	8.77E-02	0.8	up
ENSG00000176749	<i>CDK5R1</i>	17	1.81	6.11	3.37	7.88E-03	9.56E-02	0.4	up
ENSG00000129355	<i>CDKN2D</i>	19	41.47	129.25	3.12	1.05E-02	9.56E-02	0.5	up
ENSG00000153922	<i>CHD1</i>	5	18.06	62.62	3.47	7.21E-03	9.54E-02	0.6	up
ENSG00000140932	<i>CMTM2</i>	16	12.84	57.63	4.49	2.63E-03	8.77E-02	0.6	up
ENSG00000115520	<i>COQ10B</i>	2	8.43	28.91	3.43	7.58E-03	9.56E-02	0.5	up
ENSG00000184371	<i>CSF1</i>	1	5.79	21.05	3.64	3.48E-03	8.77E-02	0.5	up
ENSG00000144655	<i>CSRNP1</i>	3	30.79	186.58	6.06	2.09E-03	8.77E-02	0.7	up
ENSG00000116761	<i>CTH</i>	1	0.98	3.04	3.11	7.46E-03	9.56E-02	0.4	up
ENSG00000163739	<i>CXCL1</i>	4	17.55	72.69	4.14	4.95E-03	9.09E-02	0.6	up
ENSG00000161921	<i>CXCL16</i>	17	7.16	99.52	13.89	1.20E-02	9.86E-02	0.8	up
ENSG00000124875	<i>CXCL6</i>	4	0.44	2.67	6.11	1.06E-03	8.69E-02	0.4	up
ENSG00000115165	<i>CYTIP</i>	2	31.68	167.33	5.28	3.77E-03	8.77E-02	0.7	up
ENSG00000175197	<i>DDIT3</i>	12	41.28	176.8	4.28	8.20E-03	9.56E-02	0.6	up
ENSG00000206047	<i>DEFA1</i>	8	431.61	4383.46	10.16	2.87E-02	1.69E-01	0.8	up
ENSG00000185918	<i>DEFA3</i>	8	1707.28	19529.75	11.44	7.48E-03	9.56E-02	0.8	up
ENSG00000164816	<i>DEFA5</i>	8	0.54	5.35	9.91	1.06E-02	9.56E-02	0.6	up
ENSG00000185559	<i>DLK1*</i>	14	13.21	0.32	-41	6.86E-02	3.30E-01	0.8	down
ENSG00000128590	<i>DNAJB9</i>	7	9.14	37.01	4.05	6.19E-03	9.23E-02	0.6	up
ENSG00000079393	<i>DUSP13</i>	10	1.74	5.51	3.16	7.78E-03	9.56E-02	0.4	up
ENSG00000117298	<i>ECE1</i>	1	17.9	60.08	3.36	1.01E-02	9.56E-02	0.6	up
ENSG00000105656	<i>ELL</i>	19	15.1	59.93	3.97	1.13E-02	9.72E-02	0.6	up
ENSG00000218839	<i>FAM138C</i>	1	12.15	41.67	3.43	7.14E-03	9.54E-02	0.5	up
ENSG00000213714	<i>FAM209B</i>	20	1.03	4.03	3.93	1.03E-02	9.56E-02	0.4	up
ENSG00000120709	<i>FAM53C</i>	5	29.49	92.09	3.12	1.08E-02	9.56E-02	0.5	up
ENSG00000165355	<i>FBXO33</i>	14	7.93	34.64	4.37	5.63E-03	9.09E-02	0.6	up
ENSG00000186431	<i>FCAR</i>	19	70.02	270.14	3.86	3.24E-03	8.77E-02	0.6	up
ENSG00000145780	<i>FEM1C</i>	5	12.65	44.21	3.5	5.52E-03	9.09E-02	0.6	up
ENSG00000126262	<i>FFAR2</i>	19	9.58	44.2	4.61	1.47E-03	8.77E-02	0.6	up
ENSG00000185897	<i>FFAR3</i>	19	0.08	1.75	22.01	5.84E-03	9.09E-02	0.5	up
ENSG00000075426	<i>FOSL2</i>	2	41.34	180.53	4.37	5.07E-03	9.09E-02	0.6	up
ENSG00000171051	<i>FPR1</i>	19	145.3	796.28	5.48	7.23E-03	9.54E-02	0.7	up
ENSG00000123689	<i>G0S2</i>	1	50.82	1086.51	21.38	1.87E-03	8.77E-02	0.9	up
ENSG00000139112	<i>GABARAPL1</i>	12	40.39	201.73	4.99	4.46E-03	8.77E-02	0.7	up
ENSG00000099860	<i>GADD45B</i>	19	36.99	173.45	4.69	5.73E-03	9.09E-02	0.7	up
ENSG00000124091	<i>GCNT7</i>	20	1.1	4.5	4.08	6.09E-03	9.20E-02	0.5	up
ENSG00000165474	<i>GJB2</i>	13	0.01	2.06	209.6	1.02E-02	9.56E-02	0.5	up
ENSG00000121742	<i>GJB6</i>	13	2.67	53.3	19.93	2.34E-02	1.44E-01	0.8	up
ENSG00000135821	<i>GLUL</i>	1	162.09	595.13	3.67	1.19E-02	9.86E-02	0.6	up
ENSG00000120063	<i>GNA13</i>	17	34.24	121.07	3.54	1.00E-02	9.56E-02	0.6	up
ENSG00000140030	<i>GPR65</i>	14	14.06	65.49	4.66	4.46E-03	8.77E-02	0.6	up
ENSG00000109519	<i>GRPEL1</i>	4	20.42	74.49	3.65	8.24E-03	9.56E-02	0.6	up
ENSG00000125812	<i>GZF1</i>	20	7.63	34.53	4.53	4.38E-03	8.77E-02	0.6	up

Gene ID	Gene	Chr	FPKM MeanWT	FPKM MeanM	FPKM Fold Change	pval	padj	P	Label
ENSG00000132475	<i>H3F3B</i>	17	326.23	1053.1	3.23	1.13E-02	9.72E-02	0.6	up
ENSG00000182782	<i>HCAR2</i>	12	5.59	40.77	7.3	1.20E-03	8.69E-02	0.7	up
ENSG00000177374	<i>HIC1</i>	17	0.15	1.22	7.86	5.11E-03	9.09E-02	0.3	up
ENSG00000100644	<i>HIF1A*</i>	14	44.67	354.33	7.93	8.00E-04	8.69E-02	0.8	up
ENSG00000187837	<i>HIST1H1C</i>	6	26.94	155.03	5.76	1.10E-02	9.68E-02	0.7	up
ENSG00000180596	<i>HIST1H2BC</i>	6	6	35.72	5.95	2.52E-03	8.77E-02	0.7	up
ENSG00000090339	<i>ICAM1</i>	19	11.24	152.76	13.59	3.79E-04	8.69E-02	0.8	up
ENSG00000067064	<i>IDI1</i>	10	25.34	104.58	4.13	5.90E-03	9.09E-02	0.6	up
ENSG00000119922	<i>IFIT2</i>	10	10	45.44	4.54	9.97E-03	9.56E-02	0.6	up
ENSG00000006652	<i>IFRD1</i>	7	29.95	120.48	4.02	6.55E-03	9.38E-02	0.6	up
ENSG00000115594	<i>IL1R1</i>	2	3.8	14.43	3.8	5.18E-03	9.09E-02	0.5	up
ENSG00000169429	<i>IL8</i>	4	60.71	4380.46	72.15	1.47E-04	8.69E-02	0.9	up
ENSG00000134070	<i>IRAK2</i>	3	2.02	7.79	3.86	6.79E-03	9.54E-02	0.5	up
ENSG00000148841	<i>ITPRIP</i>	10	26.48	136.29	5.15	2.99E-03	8.77E-02	0.7	up
ENSG00000116679	<i>IVNS1ABP</i>	1	92.24	331.42	3.59	1.21E-02	9.86E-02	0.6	up
ENSG00000070495	<i>JMJD6</i>	17	32.75	115.38	3.52	1.03E-02	9.56E-02	0.6	up
ENSG00000157551	<i>KCNJ15</i>	21	4.4	29.42	6.68	1.19E-02	9.86E-02	0.7	up
ENSG00000102554	<i>KLF5</i>	13	5.17	22.01	4.25	2.92E-03	8.77E-02	0.6	up
ENSG00000174010	<i>KLHL15</i>	X	9.7	42.55	4.39	4.18E-03	8.77E-02	0.6	up
ENSG00000108244	<i>KRT23</i>	17	2.18	21.27	9.74	1.13E-02	9.72E-02	0.7	up
ENSG00000196878	<i>LAMB3</i>	1	0.11	4.78	43.93	3.27E-04	8.69E-02	0.7	up
ENSG00000189067	<i>LITAF</i>	16	77.84	323.42	4.15	8.19E-03	9.56E-02	0.6	up
ENSG00000164715	<i>LMTK2</i>	7	5.85	17.82	3.05	9.72E-03	9.56E-02	0.5	up
ENSG00000171236	<i>LRG1</i>	19	12.92	49.71	3.85	8.26E-03	9.56E-02	0.6	up
ENSG00000128594	<i>LRRC4</i>	7	5.47	28.48	5.2	2.36E-03	8.77E-02	0.6	up
ENSG00000185022	<i>MAFF</i>	22	4.35	27.29	6.27	3.84E-03	8.77E-02	0.7	up
ENSG00000140941	<i>MAP1LC3B</i>	16	33.04	129.49	3.92	1.04E-02	9.56E-02	0.6	up
ENSG00000171471	<i>MAP1LC3B2</i>	12	14.98	77.32	5.16	4.92E-03	9.09E-02	0.7	up
ENSG00000069956	<i>MAPK6</i>	15	6.62	21.4	3.23	9.47E-03	9.56E-02	0.5	up
ENSG00000155130	<i>MARCKS</i>	6	13.87	63.68	4.59	4.77E-03	9.09E-02	0.6	up
ENSG00000186594	<i>MIR22HG</i>	17	14.43	116.6	8.08	8.35E-04	8.69E-02	0.8	up
ENSG00000196549	<i>MME</i>	3	5.07	55.88	11.01	1.03E-03	8.69E-02	0.8	up
ENSG00000100985	<i>MMP9*</i>	20	110.87	673.47	6.07	1.03E-02	9.56E-02	0.7	up
ENSG00000139505	<i>MTMR6</i>	13	10.51	34.07	3.24	1.05E-02	9.56E-02	0.5	up
ENSG00000059728	<i>MXD1</i>	2	97.09	357.83	3.69	5.31E-03	9.09E-02	0.6	up
ENSG00000105835	<i>NAMPT</i>	7	39.25	615.49	15.68	1.19E-03	8.69E-02	0.9	up
ENSG00000165030	<i>NFIL3</i>	9	42.75	191.04	4.47	7.20E-03	9.54E-02	0.6	up
ENSG00000077150	<i>NFKB2</i>	10	20.85	62.65	3	9.72E-03	9.56E-02	0.5	up
ENSG00000100906	<i>NFKBIA*</i>	14	79.58	682.74	8.58	3.78E-03	8.77E-02	0.8	up
ENSG00000144802	<i>NFKBIZ</i>	3	63.28	216.45	3.42	6.24E-03	9.23E-02	0.6	up
ENSG00000131669	<i>NINJ1</i>	9	31.82	171.73	5.4	3.74E-03	8.77E-02	0.7	up
ENSG00000153234	<i>NR4A2*</i>	2	9.01	40.55	4.5	1.19E-02	9.86E-02	0.6	up
ENSG00000119508	<i>NR4A3</i>	9	0.14	12.56	88.77	1.30E-05	1.91E-02	0.8	up
ENSG00000175352	<i>NRIP3</i>	11	0.71	4.63	6.51	5.96E-03	9.09E-02	0.5	up
ENSG00000135114	<i>OASL</i>	12	4.87	14.61	3	1.16E-02	9.77E-02	0.5	up
ENSG00000099985	<i>OSM</i>	22	6	57.64	9.6	1.25E-03	8.69E-02	0.8	up
ENSG00000184588	<i>PDE4B</i>	1	6.49	73.06	11.26	2.91E-04	8.69E-02	0.8	up

Gene ID	Gene	Chr	FPKM MeanWT	FPKM MeanM	FPKM Fold Change	pval	padj	P	Label
ENSG00000165650	<i>PDZD8</i>	10	25.24	92.53	3.67	3.40E-03	8.77E-02	0.6	up
ENSG00000197329	<i>PELI1</i>	2	18.52	114.46	6.18	1.14E-03	8.69E-02	0.7	up
ENSG00000170525	<i>PFKFB3</i>	10	27.89	143.6	5.15	1.83E-03	8.77E-02	0.7	up
ENSG00000116273	<i>PHF13</i>	1	4.24	14.44	3.4	9.83E-03	9.56E-02	0.5	up
ENSG00000124102	<i>PI3</i>	20	5.28	46.26	8.77	1.23E-03	8.69E-02	0.7	up
ENSG00000165195	<i>PIGA</i>	X	6.86	30.59	4.46	3.51E-03	8.77E-02	0.6	up
ENSG00000011422	<i>PLAUR*</i>	19	24.09	196.33	8.15	3.67E-03	8.77E-02	0.8	up
ENSG00000115956	<i>PLEK</i>	2	167.69	657.66	3.92	5.72E-03	9.09E-02	0.6	up
ENSG00000173846	<i>PLK3</i>	1	11.77	63.67	5.41	4.26E-03	8.77E-02	0.7	up
ENSG00000135241	<i>PNPLA8</i>	7	11.31	45.29	4	6.86E-03	9.54E-02	0.6	up
ENSG00000158615	<i>PPP1R15B</i>	1	24.09	96.07	3.99	4.49E-03	8.77E-02	0.6	up
ENSG00000186652	<i>PRG2</i>	11	27.02	442.68	16.38	2.20E-01	8.19E-01	0.9	up
ENSG00000156575	<i>PRG3</i>	11	0.32	43.33	134.67	1.10E-01	4.84E-01	0.9	up
ENSG00000103355	<i>PRSS33</i>	16	0.02	1.62	78.7	1.13E-02	9.72E-02	0.5	up
ENSG00000073756	<i>PTGS2*</i>	1	20.01	250.94	12.54	4.96E-02	2.58E-01	0.8	up
ENSG00000112245	<i>PTP4A1</i>	6	24.87	86.5	3.48	8.75E-03	9.56E-02	0.6	up
ENSG00000132334	<i>PTPRE</i>	10	37.25	158.57	4.26	4.14E-03	8.77E-02	0.6	up
ENSG00000155093	<i>PTPRN2</i>	7	2.15	9.3	4.32	1.02E-02	9.56E-02	0.5	up
ENSG00000115828	<i>QPCT</i>	2	33.29	121.77	3.66	9.93E-03	9.56E-02	0.6	up
ENSG00000166128	<i>RAB8B</i>	15	24	79.39	3.31	1.04E-02	9.56E-02	0.6	up
ENSG00000154710	<i>RABGEF1</i>	7	10.1	44.8	4.44	5.56E-03	9.09E-02	0.6	up
ENSG00000177483	<i>RBM44</i>	2	0.63	1.99	3.16	9.02E-03	9.56E-02	0.3	up
ENSG00000162444	<i>RBP7</i>	1	2.94	13.13	4.47	8.23E-03	9.56E-02	0.6	up
ENSG00000139890	<i>REM2</i>	14	4.06	13.51	3.33	8.89E-03	9.56E-02	0.5	up
ENSG00000116741	<i>RGS2</i>	1	241.35	1403.63	5.82	2.36E-03	8.77E-02	0.7	up
ENSG00000101782	<i>RIOK3</i>	18	37.26	138.59	3.72	3.81E-03	8.77E-02	0.6	up
ENSG00000104312	<i>RIPK2</i>	8	9.24	65.52	7.09	2.21E-03	8.77E-02	0.7	up
ENSG00000143622	<i>RIT1</i>	1	19.06	88.81	4.66	4.15E-03	8.77E-02	0.6	up
ENSG00000134758	<i>RNF138</i>	18	12.32	41.84	3.4	8.20E-03	9.56E-02	0.5	up
ENSG00000116514	<i>RNF19B</i>	1	16.1	83.82	5.21	2.70E-03	8.77E-02	0.7	up
ENSG00000052749	<i>RRP12</i>	10	14.96	45.93	3.07	1.04E-02	9.56E-02	0.5	up
ENSG00000163221	<i>S100A12</i>	1	422.88	2568.17	6.07	9.32E-03	9.56E-02	0.7	up
ENSG00000163220	<i>S100A9</i>	1	5879.73	28114.96	4.78	7.08E-03	9.54E-02	0.7	up
ENSG00000163993	<i>S100P</i>	4	133.56	837.74	6.27	1.47E-03	8.77E-02	0.7	up
ENSG00000155307	<i>SAMSN1</i>	21	29.85	131.11	4.39	5.91E-03	9.09E-02	0.6	up
ENSG00000130066	<i>SAT1</i>	X	355.35	1473.87	4.15	8.13E-03	9.56E-02	0.6	up
ENSG00000129657	<i>SEC14L1</i>	17	31.33	153.62	4.9	3.73E-03	8.77E-02	0.7	up
ENSG00000130766	<i>SESN2</i>	1	4.82	28.3	5.88	3.91E-03	8.77E-02	0.7	up
ENSG00000175793	<i>SFN</i>	1	0.92	4.18	4.57	8.92E-03	9.56E-02	0.5	up
ENSG00000118515	<i>SGK1</i>	6	21.51	97.98	4.55	1.04E-02	9.56E-02	0.6	up
ENSG00000142178	<i>SIK1</i>	21	5.56	57.13	10.27	1.33E-02	1.01E-01	0.8	up
ENSG00000116991	<i>SIPA1L2</i>	1	1.95	7.97	4.08	3.52E-03	8.77E-02	0.5	up
ENSG00000163877	<i>SNIP1</i>	1	8.58	28.73	3.35	1.04E-02	9.56E-02	0.5	up
ENSG00000184602	<i>SNN</i>	16	18.14	74.73	4.12	5.70E-03	9.09E-02	0.6	up
ENSG00000163788	<i>SNRK</i>	3	27.42	88.53	3.23	1.23E-02	9.92E-02	0.5	up
ENSG00000178996	<i>SNX18</i>	5	9.73	37.82	3.89	5.02E-03	9.09E-02	0.6	up
ENSG00000112096	<i>SOD2</i>	6	107.83	589.85	5.47	4.13E-03	8.77E-02	0.7	up

Gene ID	Gene	Chr	FPKM MeanWT	FPKM MeanM	FPKM Fold Change	pval	padj	P	Label
ENSG00000118785	SPP1	4	0.16	4.28	26.23	1.16E-02	9.77E-02	0.6	up
ENSG00000136158	SPRY2	13	2.31	10.06	4.37	7.09E-03	9.54E-02	0.5	up
ENSG00000179119	SPTY2D1	11	7.35	34.37	4.67	3.63E-03	8.77E-02	0.6	up
ENSG00000081320	STK17B	2	68.88	352.2	5.11	2.35E-03	8.77E-02	0.7	up
ENSG00000169895	SYAP1	X	10.68	68.76	6.44	6.53E-03	9.38E-02	0.7	up
ENSG00000197780	TAF13	1	12.05	71.05	5.9	1.78E-03	8.77E-02	0.7	up
ENSG00000134827	TCN1	11	6.8	72.55	10.67	9.46E-03	9.56E-02	0.8	up
ENSG00000140332	TLE3	15	21.32	87.53	4.11	5.43E-03	9.09E-02	0.6	up
ENSG00000057704	TMCC3	12	3	27.32	9.12	3.95E-03	8.77E-02	0.7	up
ENSG00000183160	TMEM119	12	0.23	1.4	6.1	1.52E-03	8.77E-02	0.3	up
ENSG00000118503	TNFAIP3	6	26.36	176.55	6.7	5.64E-03	9.09E-02	0.7	up
ENSG00000123610	TNFAIP6	2	6.07	21.07	3.47	1.24E-02	9.92E-02	0.5	up
ENSG00000006327	TNFRSF12A*	16	1.66	7.38	4.45	2.96E-03	8.77E-02	0.5	up
ENSG00000078804	TP53INP2	20	11.12	62.43	5.62	1.80E-03	8.77E-02	0.7	up
ENSG00000124731	TREM1	6	48.92	439.35	8.98	1.02E-03	8.69E-02	0.8	up
ENSG00000112195	TREML2	6	8.71	35.4	4.06	3.19E-03	8.77E-02	0.6	up
ENSG00000188056	TREML4	6	1.05	9.69	9.19	1.49E-03	8.77E-02	0.6	up
ENSG00000196428	TSC22D2	3	7.27	24.12	3.32	8.96E-03	9.56E-02	0.5	up
ENSG00000175894	TSPEAR	21	0.12	4.31	34.7	4.28E-04	8.69E-02	0.6	up
ENSG00000165006	UBAP1	9	24.37	76.75	3.15	1.24E-02	9.92E-02	0.5	up
ENSG00000162543	UBXN10	1	0.9	3.05	3.37	1.01E-02	9.56E-02	0.4	up
ENSG00000147852	VLDLR	9	0.81	3.57	4.4	1.79E-03	8.77E-02	0.4	up
ENSG00000155659	VSIG4	X	41.95	1.8	-23.31	2.83E-01	9.59E-01	0.8	down
ENSG00000156232	WHAMM	15	6.38	19.39	3.04	7.75E-03	9.56E-02	0.5	up
ENSG00000119801	YPEL5	2	59.82	216.66	3.62	7.68E-03	9.56E-02	0.6	up
ENSG00000169155	ZBTB43	9	6.27	19.74	3.15	8.75E-03	9.56E-02	0.5	up
ENSG00000204160	ZDHC18	1	39.1	152.33	3.9	4.23E-03	8.77E-02	0.6	up
ENSG00000141497	ZMYND15	17	0.86	6.24	7.22	7.60E-03	9.56E-02	0.6	up
ENSG00000185947	ZNF267	16	11.12	34.42	3.1	1.07E-02	9.56E-02	0.5	up
ENSG00000162702	ZNF281	1	13.67	55.6	4.07	3.60E-03	8.77E-02	0.6	up
ENSG00000173276	ZNF295	21	2.92	9.56	3.27	9.36E-03	9.56E-02	0.5	up

Table 8. DEG in SETBP1-mutated samples compared to non-mutated aCML samples. For each listed gene, FPKM mean values under each condition, corresponding FPKM Fold Change, adjusted p values (DESeq) and P probability (NOISeq) are reported, in addition to a flag (up/down) indicating the relative expression level in our aCML sample set. Genes related to TGFB are labeled by *. (R Piazza et al. Recurrent SETBP1 mutations in atypical chronic myeloid leukemia. Nat Genet 2013;45:18-24)

Immunofluorescence. 293T cells expressing wild-type SETBP1 or SETBP1 Gly870Ser were seeded on glass coverslips in a 6-well plate. After 24 h, cells were washed, fixed using 4% paraformaldehyde and were incubated at room temperature for 30 minutes. Subsequently, cells were treated with a buffer composed of PBS (pH 7.4) containing 0.1 M glycine and 0.3% Triton X-100 and stained with Alexa Fluor 546-conjugated phalloidin (Invitrogen) for 1 h at room temperature to define the cellular morphology, and with TOTO-3 iodide (Invitrogen) to obtain nuclear definition. Confocal microscopy was performed using a Radiance 2100 laser scanning confocal microscope (Bio-Rad) having a krypton/argon laser and a red laser diode.

Results

Exome sequencing of aCML samples

Next-generation sequencing (NGS) technologies have revolutionized and transformed the genetics study of human diseases. Exome sequencing, that consists in the targeted sequencing of the protein-coding portion of the human genome, in particular, has been shown to be a powerful and cost-effective method for detection of disease variants (89).

Considering our lack of knowledge about aCML somatic lesions and because no specific recurrent genomic or karyotypic abnormalities have been yet identified, we decided to apply a high-throughput strategy to aCML, including both exome-sequencing and RNA-sequencing, in order to identify recurrent mutations. In particular, we applied exome sequencing to eight individuals with aCML, by comparing DNA from tumoral leukocytes and constitutive DNA extracted from lymphocytes (81), as specified in material and methods. Each read of a massively parallel sequencing run is clonal and therefore derives from a single molecule of genomic DNA (81). In this way, the proportion of sequencing reads showing a variant allele can be quantitatively correlated to the proportion of cells in the DNA sample carrying the mutation, considering an adequate coverage for the gene under analysis. As the NGS technologies, in general, including the exome-sequencing, use PCR-like amplification step in the library preparation, multiple reads originating from the same template can be sequenced, and overrepresentation of certain alleles due to amplification bias can interfere in the variant calling step. For this reason, to remove PCR duplicates and to minimize the detection of sub-clonal variations, we decided to apply specific algorithms dedicated to the removal of PCR duplicates and we considered only mutations with a frequency of at least 35%. This means that, considering that the minimum relative coverage of each variant in the

filtering algorithms (see materials and methods) was set to 0.35 and that the percentage of leukemic cells in our experiments is >80% and hypothesizing that the mutations are present as heterozygous variations, when the mutation occurs in 35% of reads, it is present in 70-87.5% of leukemic cells. By filtering data using this minimum relative coverage, the exome sequencing analysis in these eight aCML patients, allowed us to identify 84 somatic mutations, including 63 nonsynonymous and 21 synonymous variations, as reported in Table 9. All the somatic mutations identified using Next-Generation Sequencing were validated with the traditional Sanger sequencing, using PCR primer pairs specific for each variant, as reported in the Table 10 and obtaining a validation rate of 96%.

GENE	GENE DESCRIPTION	MUTATION	LOCUS	MUTATION FRACTION (%)	ABSOLUTE COVERAGE	MUTATION FRACTION IN CONTROL (%)	ABSOLUTE COVERAGE IN CONTROL	PATIENT #
<i>GNE</i>	Bifunctional UDP-N-acetylglucosamine 2-epimerase/N-acetylmannosamine kinase	Pro704Ser	9,36207420-36207421,1,G/A	54.9	51	0	89	1
<i>HOXC8</i>	Homeobox protein Hox-C8	Arg27Gln	12,52689414-52689415,1,G/A	39.3	89	0.9	213	1
<i>IDH2</i>	Isocitrate dehydrogenase	Arg140Gln	15,88432937-88432938,1,C/T	59.2	49	0	101	1
<i>PTPRT</i>	Receptor-type tyrosine-protein phosphatase T Precursor	Met1287Ile	20,40160486-40160487,1,C/T	39.6	149	0.8	260	1
<i>RIT1</i>	GTP-binding protein Rit1	Asp68Val	1,154146875-154146876,1,T/A	44.3	88	0	282	1
<i>SFRS2</i>	Splicing Factor, Arginine/Serine-rich, 2	Pro95Leu	17,72244553-72244554,1,G/A	41.9	31	0	46	1
<i>SI</i>	Sucrase-isomaltase, intestinal	Arg1484His	3,166199110-166199111,1,C/T	42.9	70	0	95	1
<i>TMEM201</i>	Transmembrane protein 201	Pro252Arg	1,9583897-9583898,1,C/G	40.8	49	0	64	1
<i>UBE2D1</i>	Ubiquitin-conjugating enzyme E2 D1	Ile106Val	10,59797699-59797700,1,A/G	39.0	41	0	71	1
<i>GIGYF2</i>	PERQ amino acid-rich with GYF domain-containing protein 2	Ser273Cys	2,233363756-233363757,1,C/G	50.8	122	0	96	2
<i>GLG1</i>	Golgi apparatus protein 1 Precursor	Ile499Thr	16,73077269-73077270,1,A/G	48.7	185	0.7	138	2
<i>MTA2</i>	Metastasis-associated protein MTA2	Tyr283Cys	11,62120718-62120719,1,T/C	60.8	232	0	131	2
<i>MYCL1</i>	L-myc-1 proto-oncogene protein	Cys98*	1,40139489-40139490,1,G/T	45.0	20	0	24	2
<i>PCDHB11</i>	Protocadherin beta-11 Precursor	Arg62Trp	5,140559714-140559715,1,C/T	35.5	76	0	51	2
<i>PTH2</i>	Parathyroid hormone 2 receptor Precursor	Lis425*	2,209066248-209066249,1,A/T	52.4	84	0	98	2
<i>SMOC2</i>	SPARC-related modular calcium-binding protein 2 Precursor	Ser225Leu	6,168741425-168741426,1,C/T	40.0	30	0	18	2
<i>CD22</i>	B-cell receptor CD22 Precursor	Val475Ile	19,40523796-40523797,1,G/A	42.9	49	0	43	3
<i>MLH3</i>	DNA mismatch repair protein Mlh3	Asn781Ser	14,74583769-74583770,1,T/C	58.3	36	1.5	68	3
<i>SETBP1</i>	SET-binding protein	Gly870Ser	18,40785910-40785911,1,G/A	52.6	38	0	35	3

<i>SIX4</i>	Homeobox protein SIX4	Gly482Ser	14,60256335-60256336,1,C/T	45.7	94	0	105	3
<i>SYNE1</i>	Nesprin-1(Nuclear envelope spectrin repeat protein 1)	Arg3768Gln	6,152715131-152715132,1,C/T	43.2	44	1.6	64	3
<i>CD163L1</i>	Scavenger receptor cysteine-rich type 1 protein M160 Precursor	Ser651Phe	12,7440055-7440056,1,G/A	38.5	78	1.3	80	4
<i>DHRS7C</i>	Dehydrogenase/reductase SDR family member 7C Precursor	Val235Met	17,9616880-9616881,1,C/T	51.6	93	1.4	69	4
<i>EZH2</i>	Histone-lysine N-methyltransferase EZH2	Asn693Lis	7,148137365-148137366,1,A/T	42.7	82	0	74	4
<i>IFNAR1</i>	Interferon-alpha/beta receptor alpha chain Precursor (IFN-alpha-REC)	Glu132Gly	21,33637461-33637462,1,A/G	41.2	148	0	102	4
<i>IRAK4</i>	Interleukin-1 receptor-associated kinase 4	Asp329His	12,42462419-42462420,1,G/C	46.2	91	0	73	4
<i>NRAS</i>	GTPase NRas Precursor (Transforming protein N-Ras)	Gly12Cys	1,115060270-115060271,1,C/A	39.2	125	0	111	4
<i>PRR21</i>	Putative proline-rich protein 21	His79Arg	2,240630836-240630837,1,T/C	84.0	25	0	5	4
<i>RBBP4</i>	Histone-binding protein RBBP4 (Retinoblastoma-binding protein 4)	Glu330Lis	1,32910658-32910659,1,G/A	44.0	91	0	97	4
<i>ACTR1A</i>	Alpha-centractin (Centractin)(Centrosome-associated actin homolog)	Arg188Trp	10,104234001-104234002,1,G/A	50.5	97	0.9	108	5
<i>ANGEL2</i>	Protein angel homolog 2	Asp108Asn	1,211245302-211245303,1,C/T	47.1	223	0	262	5
<i>ASXL1</i>	Putative Polycomb group protein ASXL1	Gln708*	20,30486297-30486298,1,C/T	59.3	81	0.8	130	5
<i>CBL</i>	E3 ubiquitin-protein ligase CBL	Cys416Ser	11,118654447-118654448,1,T/A	42.6	47	0	74	5
<i>DCN</i>	Decorin Precursor (Bone proteoglycan II)	Ala109Thr	12,90076416-90076417,1,C/T	48.2	56	0	50	5
<i>EPS15L1</i>	Epidermal growth factor receptor substrate 15-like 1	Arg430Gln	19,16376537-16376538,1,C/T	59.7	57	0	83	5
<i>GABRA5</i>	Gamma-aminobutyric acid receptor subunit alpha-5 Precursor	Phe433Leu	15,24776035-24776036,1,C/G	48.1	79	0	93	5
<i>MAGEA1</i>	Melanoma-associated antigen 1 (MAGE-1)	Tyr232His	X,152135510-152135511,1,A/G	98.1	104	2.1	144	5

	antigen)								
<i>PON2</i>	Serum paraoxonase/arylesterase 2	Pro295Leu	7,94873388-94873389,1,G/A	57.3	89	1.6	123	5	
<i>RASEF</i>	RAS and EF-hand domain-containing protein (Ras-related protein Rab-45)	Arg659Cys	9,84797705-84797706,1,G/A	48.9	180	0	234	5	
<i>SETBP1</i>	SET-binding protein	Gly870Ser	18,40785910-40785911,1,G/A	47.2	72	0	90	5	
<i>TNMD</i>	Tenomodulin (TeM)(hTeM)(Chondromodulin-I-like protein)	Arg233His	X,99740788-99740789,1,G/A	97.4	77	0	113	5	
<i>TUBA1C</i>	Tubulin alpha-1C chain (Tubulin alpha-6 chain)	Cys4Trp	12,47949522-47949523,1,C/G	43.0	135	0	226	5	
<i>UNKL</i>	RING finger protein unkempt-like	Ala127Thr	16,1393254-1393255,1,C/T	52.0	77	0	94	5	
<i>ANXA10</i>	Annexin A10 (Annexin-10)(Annexin-14)	Ile285Thr	4,169342354-169342355,1,T/C	50.7	229	0.4	285	6	
<i>DNAH9</i>	Dynein heavy chain 9, axonemal	Arg4106Pro	17,11777940-11777941,1,G/C	50.4	115	0.5	204	6	
<i>KIAA1217</i>	Sickle tail protein homolog	Ala807Val	10,24853460-24853461,1,C/T	50.0	22	0	23	6	
<i>LRP1B</i>	Low-density lipoprotein receptor-related protein 1B Precursor	Thr4441Pro	2,140721127-140721128,1,T/G	45.0	149	0	128	6	
<i>PAICS</i>	Multifunctional protein ADE2	Asp247Ala	4,57009720-57009721,1,A/C	47.4	192	0	219	6	
<i>SIGLECP3</i>	HSPC078 Fragment	Asp32Asn	19,56362620-56362621,1,G/A	45.7	140	0	188	6	
<i>UTRN</i>	Utrophin (Dystrophin-related protein 1)	His1474Tyr	6,144853913-144853914,1,C/T	51.4	70	0	68	6	
<i>VANGL2</i>	Vang-like protein 2 (Van Gogh-like protein 2)	Tyr200Cys	1,158655821-158655822,1,A/G	53.1	49	0	98	6	
<i>CASKIN2</i>	Caskin-2	Ser1042Asn	17,71009624-71009625,1,C/T	66.7	21	0	18	7	
<i>GATA2</i>	Endothelial transcription factor GATA-2 (GATA-binding protein 2)	Lis390Glu	3,129682826-129682827,1,T/C	51.5	171	0.7	141	7	
<i>GPT</i>	Alanine aminotransferase 1 (ALT1)	His14Gln	8,145700536-145700537,1,T/G	50.0	44	4.8	21	7	
<i>PSG4</i>	Pregnancy specific beta-1-glycoprotein 4	Arg320His	19,48391015-48391016,1,C/T	47.5	198	0	205	7	
<i>STOX2</i>	Storkhead-box protein 2	Arg783Cys	4,185169259-185169260,1,C/T	60.7	28	0	29	7	

<i>U2AF1</i>	Splicing factor U2AF 35 kDa subunit	Arg156His	21,43387848-43387849,1,C/T	44.3	113	1	97	7
<i>ZNF621</i>	Zinc finger protein 621	Pro342Ser	3,40549288-40549289,1,C/T	41.3	63	0	53	7
<i>ETNK1</i>	Ethanolamine kinase 1	His243Tyr	12,22703257-22703258,1,C/T	46.7	92	0	44	8
<i>EZH2</i>	Histone-lysine N-methyltransferase EZH2	Asp664Gly	7,148138395-148138396,1,T/C	45.1	142	1	102	8
<i>GPR171</i>	G-protein coupled receptor 171 (G-protein coupled receptor H963)	Gln250*	3,152399115-152399116,1,G/A	45.2	208	0	166	8
<i>NEK4</i>	Serine/threonine-protein kinase Nek4	Gln623His	3,52753319-52753320,1,C/G	51.6	161	0	110	8
<i>WHSC1L1</i>	Histone-lysine N-methyltransferase NSD3	Ser800Cys	8,38281963-38281964,1,G/C	44.9	89	0	70	8

Table 9. Somatic mutations identified in eight aCML patients by whole exome sequencing. * indicates a stop codon. The “absolute coverage” column indicates the coverage obtained for the nucleotide where the mutation was identified. “Fraction” indicates the proportion of total reads that were mutated. (R Piazza et al. Recurrent SETBP1 mutations in atypical chronic myeloid leukemia. Nat Genet 2013;45:18-24)

GENE	FORWARD PRIMERS (5'--3')	REVERSE PRIMERS (5'--3')
ACTR1A	GTAGCCATCACCCAAGTGAAG	GAACTTGGAGAAGATCAGCAAC
	GAAACAAAACCAAACCTATG	TATTTCTTCCCAGCGTGAGG
ANGEL2	CTCGCCTTGGATTATACAACAG	GGCAACGGGAGACTGATAATAC
	GCAGAGCAACTATGTCCACTC	CCTGTTAGACCCAGTGTCAAAC
ANXA10	CCAGTGTTGGAAGAAGTCAAGG	ATCTCCTGACCTCGTGATCTG
	CTGGAATAGAGAGCACAAGGC	CACCTGCCGATGTTCTAAAGA
ASXL1	GAGAGGTCACCACTGCCATAG	CTTTCTAATGTATCATCTCCCAC
	CAGCCCACTTACCAGATATGC	ACTGAGGCTGCTCCACTAATC
CASKIN2	GTTCTGTCTCCCCATTCCAC	GGAACCTGACCATCAAACAG
	CACAGTCCCAGTGCTCTCAG	GAATCAGACACTGTAAAGCGG
CBL	CTATGTAATAACCTTGAAAATTCG	GTAAAGTGCTTAACACAATATCTGTC
	TCCAACATATGTAATAATATGTGCTG	GTGTTTTACGGCTTTAGAAGACA
CD163L1	GTATGTTAGAGAACCCAGGCAG	CACACCCATTAACACTTGTCTTC
	GATCCCATAGAAAGATATTAATAATG	GGAAGACTGGAGGTGTACTTTC
CD22	CAGATCCTCCAAGAAGGTG	GCTGACCGAGTTTCCAGAGTG
	GGACAGGTAGCGGGAGAAGA	TTGCCTTCATCCCCTCATCC

<i>DCN</i>	GCACTTTGGTGATCTCATTCT	GATCCTCCTGTCTCAGCCTC
	CATCACTGTAAGGTCCAGATTC	TCAAGTGATCCTCCACCTC
<i>DHRS7C</i>	GAGTACCCATCTGAAGGTCATT	CCCACAATCACATCACACCA
	CCAGAGATAGGCAGTTGCTTG	CAGATTCTTGAGGATCTTCCC
<i>DNAH9</i>	GCTCTCAATCTCAAGAAGCG	GTCTAAGCCCTTTTGCCTCAG
	CCAAAGTGCTGGAATTACAGG	CAGAGCAGAGTCTCACCTGATG
<i>EPS15L1</i>	TGACCCAGTAGGATGTCTAAGTG	GTTTCCCTGTGACCCCCTTG
	GTGTGTTCAAGGCACAAGTTC	ATCGGGAAATACCTGCTGAG
<i>ETNK1</i>	GTGGGGATACTCTTGCTAAACTG	TCTGCCCTTGCTGTTGTAG
	CATCAAAGTAAGGGGTTTCTGG	GCTATCTGACAGCCAAAGAATC
<i>EZH2</i>	CTGGTGTCAGTGAGCATGAAG	CCTATTCGTGATGTTTGAAG
	GCAAAGTGACCCATCAAAG	CAAACCCTGAAGAACTGTAACC
<i>EZH2</i>	CAGTTCCTTTCAAGCAAGCAG	GGTTCCTCACTTCTCCAAGG
	ACTGGAGGAGGTGTGCTGAG	GAAAGAGAACTTGCTGTAGTGAC
<i>GABRA5</i>	CAACTGGGAAGATGTCTCACC	CGTATCCTGTTTTCCAGAAAG
	CTTTAGAACTCCAAGGTGAGC	GGGACATCTTGAAACATCTG
<i>GATA2</i>	GGAGTGGGCAGGATGTGTC	GGTCTCGGACTAGGGAAGTG
	CTGCTAAGGGTTTGGTCCAC	GTGATGGGACTATGAAGGTGCG
<i>GIGYF2</i>	GTTCTCTTTACTAGGTGAGTCC	CCTACTCTATCTGTTTTCTCTCC
	GTGCGTACTTGGAATACAGTG	GTCCATCTCCTGCTTCTCTG
<i>GLG1</i>	GTTCCCTGTGCTGCCAAATG	CATGTTGTAAGACCTCCAGG
	CCTATCCTGAGTGTTTCTCTCT	GGTAAATAGGACTGAGCAACTG
<i>GNE</i>	AAGAAACGGTCATCC	TTCTGCCTCCTCTTTGGTATTTTG
	AGGTGACTCTGGAAGAGGAGAC	CATACCATGAATCCCTCCCTTG
<i>GPR17</i>	GGTGAGGCAAAAGTCTCAGTG	GTCAAATGTGGGTTGTATGGAG
	GCTTGACTTGCAATTTAGAAACAG	GAAGAAATTGGCATTGCTGAC

<i>GPT</i>	TGCTCTCTGCCTGGAGTTC	TGAAAGGCTTCTTCACACCC
	CATGTGGGTCTCCTGCTCTC	TGTGCCCTACCTAACCCCTC
<i>HOXC8</i>	AGCTCCTACTTCGTCAACCC	TCGCTACTGTTAGTGTGGCGG
	TACTCGTGAGCCAGAGGGGAG	TGCTGGTAGCCTGAGTTGGAG
<i>IDH2</i>	TACTTCCCCTCCTTGACAC	TGGTTGAAAGATGGCGGCTGC
	CAAGCCAGCCTCACCTCGTC	GGACCACTATTATCTCTGCCTC
<i>IFNAR1</i>	CAGTGTTAGACTGCTGCTGG	CCTGTGGTCTTTATAACAATGTACTG
	CAGTCTCTTCTGGCTTCATC	TTCAATCCTTTCCTATAACACAA
<i>IRAK4</i>	CGTTACTCTGTAAATATGTATG	ATAAGAGAGATAAACAAGATGGGTG
	GAGGACAGTTGCTTCTTTAGC	CATTTTCACAATTATAGCAGAATG
<i>KIAA1217</i>	CACTGATGGGCTCCTGAAAG	CCATACACAATCATGTCATCATC
	CACCAGCAGAAAGGGATATG	GACTCCTTTATGGGGTTATGAG
<i>LRP1B</i>	CCATGAAGAAAGTGTCTAAGACTC	TTCTTAATCAAGATTTCAATATCGC
	CCTGGAGCATTTTCACATCC	GCTCCTGTAACCTTTCATTTGG
<i>MAGEA1</i>	TCATGCTCAGACTCCCTCTTC	GGCAGAAATGCTGGAGAGTG
	CAAGTCACCCAATAGAACAGAAAC	GATAATCAGATCATGCCCAAGAC
<i>MLH3</i>	TCTCTGAGTTAAGGATGTGGC	GCTAGAGAATGAACCTACAGC
	GAGGCTAGTGATTCAGATGAC	CAAACCAATCGTCCGTAAGAAG
<i>MTA2</i>	GGCAGAGGGAGGAAGAATGA	GCCAGGATGATCTCGATCTC
	CACACTTACTAGGTGGGAATG	ATCTTCCCATAGTTTCACGC
<i>MYCL1</i>	TATGATGGAGGCGTAGTTCC	GTTGGGGAGGAACGAGAGC
	GGTCCTTACCCGAGTCGCTT	GAGGATTTCTACCGCTCCAC
<i>NEK4</i>	CCTCCTCTGATATTTGAACCAC	AAGCAATTCTCCTGCCTCAG
	GGTCTGAAACCTCCCAGGAC	CAGAGTGCTGGGATTACAAGTG
<i>NRAS</i>	GATGATCCGACAAGTGAGAGAC	GAGACAGAAGGGAGAATGGG
	ACAGAACTGTTAAAACTGGACAG	CCTTCGGGGAGTAATAGGAAG

<i>PAICS</i>	GAAGGTACAGATAAAACAAGTACAGA	GAGACTTGAATAGAGTCCCAGA
	CACAGTGGTCTGAGGAACAGC	CACAAAACAGGTTTCATATCATTAG
<i>PCDHB11</i>	GAACCAAGGGACACGCACTC	GTTCTCTGGGATTTCTAGGAGCA
	CACTGGATGGCGTTGTGAAC	GCTCAATTTGTAAAACTGCGTG
<i>PON2</i>	CAATTACAAAACCACGACATAAG	CAGAGACCGATTATCACCCCTAAC
	GTGCCTACTTTCACACTACTTCAAC	AGATTTTGACAATTTCTCATTTG
<i>PRR21</i>	GTGAAGAGGCATGGAGGAAG	TGTTTCATCCACAGCTCTTCATC
	GTGAAGAGGCATGGACGAAG	GCTTTTCATTCACTTTTCTTC
<i>PTHR2</i>	CATAGATAAAGTCCTGGCAC	CCTTGGCATCCTTCAGTGTC
	CTGAAGTTCTCACTGGCAATG	TCCAGACATAGCCAGGTAAG
<i>PTPRT</i>	GCCAGACACTGCTTATTACAAGAC	GCCTCATCATCATCAGAATATCACTG
	GGTGACAGGACCCTCTGTAGG	CCTCTACCCAACACCGTGGC
<i>RASEF</i>	CTCAATGGGAGAAAATGATTATG	GTTTCCTTTGCCATAAACTGA
	GTGAGAGGGAGGAGGAAGATG	ATGTATGGCTGTTTTAGAGATACTT
<i>RBBP4</i>	CTTGAACCTCTGACCTCATGTG	TGCAGATAGAGGACATTAATTAC
	CTCTGCTTCTGGGTTCAAG	GCTTTTAACATACCAACAACCTG
<i>RIT1</i>	GTTGCTTACCAACTGCTGATACC	CCAGATGGGATACTTGCTATCTTG
	AAGATGGCTTATGACATAAAGTCTG	ATTCTCTCTGTGGGTCACCTTTG
<i>SFRS2</i>	GTACCTGCGGGGTGGCGGTC	GGGTATGACCTCCCTCAAGGTG
	CCCGTTTACCTGCGGCTCCG	GACCGCTACACCAAGGAGTC
<i>SI</i>	GTCCATGTTGTCCCATCGTGCA	GCATGTCACTCTGGTAAATGTTGG
	CACCTCAATCCAGCTTCTG	GAGAGAAGACAACTTATCAAGAGC
<i>SIGLECP3</i>	CTCTGCTCACACAGGAAGCC	GCCTTACCTGTACATACACAG
	CAGTCTCCCTCCACTCTGTG	CCAAAAGAAGTATGAACCGTTG
<i>SIX4</i>	CACCAAATGCGTCACTTACAG	TTTTATTCAGGGACCCAGTGG
	GCAAGGGAATAAAGCAACCTG	AGTCCTCCAGAGTTCTGCTAA

<i>SMOC2</i>	TGACTTGTGAGCTGTGCTGC	TGTGTGGGTATGTGTGTGGTA
	CTCGTCACATTCAGGAGTCC	GACTGTGTGTCGGTGTGTGT
<i>STOX2</i>	GGACAAGAGGAAAGAGATATTTAGC	GAGACACCTCTCCCTACCGTG
	CACTCTGACTTTGGCAGAAGG	CTCTGGTTGCTGCTGCTGTC
<i>SYNE1</i>	GGTATTCTGCTATCCACTGTG	AGAAGTTCAAGACCAGCCTGA
	AGAATCACCTCCCTCCCATTT	CTAGCCTGGGTGACAAAGTG
<i>TMEM201</i>	ATCCTGCTCCGTGCCCTCG	CCCTAGACAGTAATCCAGCACACA
	GTGTGAGAACATGGGTCTGGC	ACACAGCTTCTCCGCCATGTG
<i>TNMD</i>	GCACTTCTTTCAGGAACATTTG	GGTTCACTAGCAAGAGTTCTCAC
	CCAGTTAGAAGACACAATGAAAATC	CCAATGATTGTGATCTGAGGG
<i>TUBA1C</i>	CTGGGATTACAGGTGTGAGTC	AGTGCGAACTTCATCTGGAG
	CTCCACCTCCTGAGTTCAAG	GAGGTCAACTACCAATGACTG
<i>U2AF1</i>	TGGGTTGGAAGGAGACATTTAC	TCCACCTTACTATACCTGAGTGTG
	ACAGAAACGACACCTATGCC	TGAGAAAAGTCTTATTAAGCGTGG
<i>UBE2D1</i>	ATAAGTACAGTATCTATATTCCTACTAAAA	GGCATACTATATAGGAATGTATC
	GAGTGTATGAGCACAGTATCC	TTTATAGATTTGTGCAATATCTGGTAC
<i>UNKL</i>	GCAAAAAAAGGACTAACATCCAG	TCTTACTGTGTTTGACTGGTTGAC
	GAATGGTGTGAACCCAGGAG	CCATGATAACCAGGCATTGC
<i>UTRN</i>	GATAATCATCCTGATGGGAATTG	GCTTATTAGTCACACTCTGCTGAAG
	CTTCTGAGATTGAGACACTGAGTG	CTTCACATGGTGATTAAGCTGTAT
<i>VANGL2</i>	ACTGCTGTCTTCTCACGC	CTATCAAGCAGCAAACATTGG
	GAAACGACGACAGTAGTAACGG	CCTAGCAGTCACTCCACATCC
<i>WHSC1L1</i>	CAGAGACGGTGGTATCCAATATG	GAATCAGTTTTGGAGATCTAAATTG
	CCTAGTCTGGCTCTGTCATCC	GCTGTTTGATGTCTGTAGCTGC
<i>ZNF621</i>	GGTCACTTTTTATTGTCCATCAG	GCATAAGAGCCGTAAGACTGC
	CCCTATGAATGTAAAGAGTGTGG	GGTATTCCAGATGTAGGCAGC

Table 10. PCR and sequencing primers used to validate the somatic mutations identified by whole exome sequencing. (R Piazza et al. Recurrent SETBP1 mutations in atypical chronic myeloid leukemia. Nat Genet 2013;45:18-24)

The analysis of transitions and transversions in the nonsynonymous alterations showed that transitions were present in 73% of the cases and transversions in 27%, as shown in Figure 8.

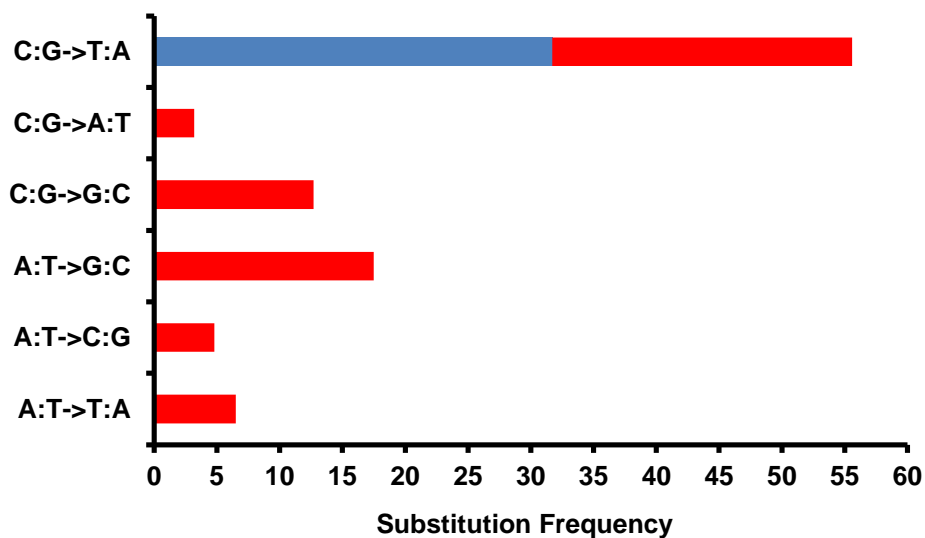


Figure 8. Distribution of substitutions in the 6 different classes of mutations. The red bars indicate the frequency of substitutions in the 8 aCML samples for the six possible mutation classes. The blue bar

indicates the proportion of C:G>T:A mutations occurring in the first base of a CpG dinucleotide. (R Piazza et al. Recurrent SETBP1 mutations in atypical chronic myeloid leukemia. Nat Genet 2013;45:18-24)

The median absolute coverage of our variant positions was 84x, ranging from 20x, relative to the *MYCL1* gene, to 232x, relative to the *MTA2* gene. Four of the alterations identified resulted in a premature stop codon: among them, a mutation in the *ASXL1* gene, that recent reports have shown mutated in several myeloid malignancies, in particular in myelodysplastic syndromes and myelomonocytic leukemia patients (90) has been identified. Another gene that we found mutated in our analysis was *IDH2* (p.Arg140Gln); the same *IDH2* alteration has been identified in Acute Myeloid Leukemia patients too, using next-generation sequencing (91). It is known that *IDH2* codes for a NADP-dependent isocitrate dehydrogenase enzyme and it is reported that mutations of *IDH2* (involving Arg in codon 140, in our case, or 172 in other cases) prevent oxidative decarboxylation of isocitrate to α -ketoglutarate and confer novel enzymatic activity, facilitating the reduction of α -ketoglutarate to d-2-hydroxyglutarate (2-)HG, a putative oncometabolite (92). To address if accumulation of the oncometabolite 2-HG was detectable in subject 1, where the somatic *IDH2* mutation (p.Arg140Gln) was identified, we performed dedicated mass spectrometry analyses in order to measure its intracellular levels. This analysis allowed us to demonstrate that intracellular 2-HG was over 10 times higher than in the normal cells used as control, or in other aCML samples, as shown in Fig 9.

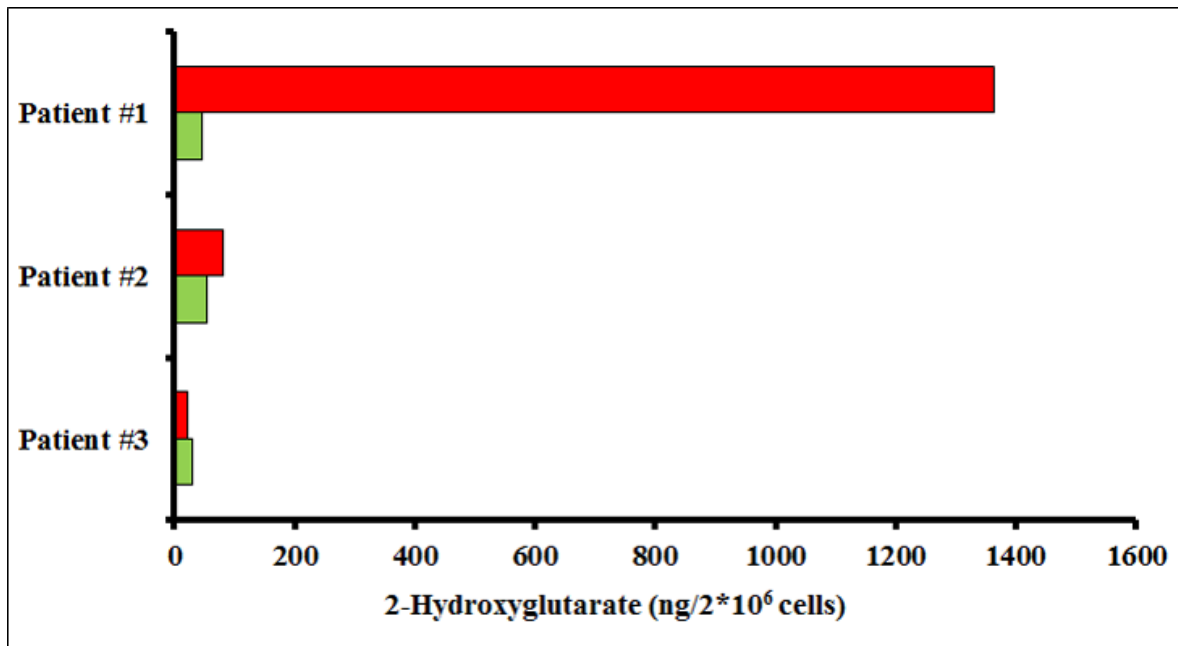


Figure 9. 2-Hydroxyglutarate levels in leukemic and normal cells derived from patient #1 in whom an IDH2 mutation was identified by Exome sequencing and in patient #2 and #3 who tested negative for IDH1/2 mutations. (R Piazza et al. Recurrent SETBP1 mutations in atypical chronic myeloid leukemia. *Nat Genet* 2013;45:18-24)

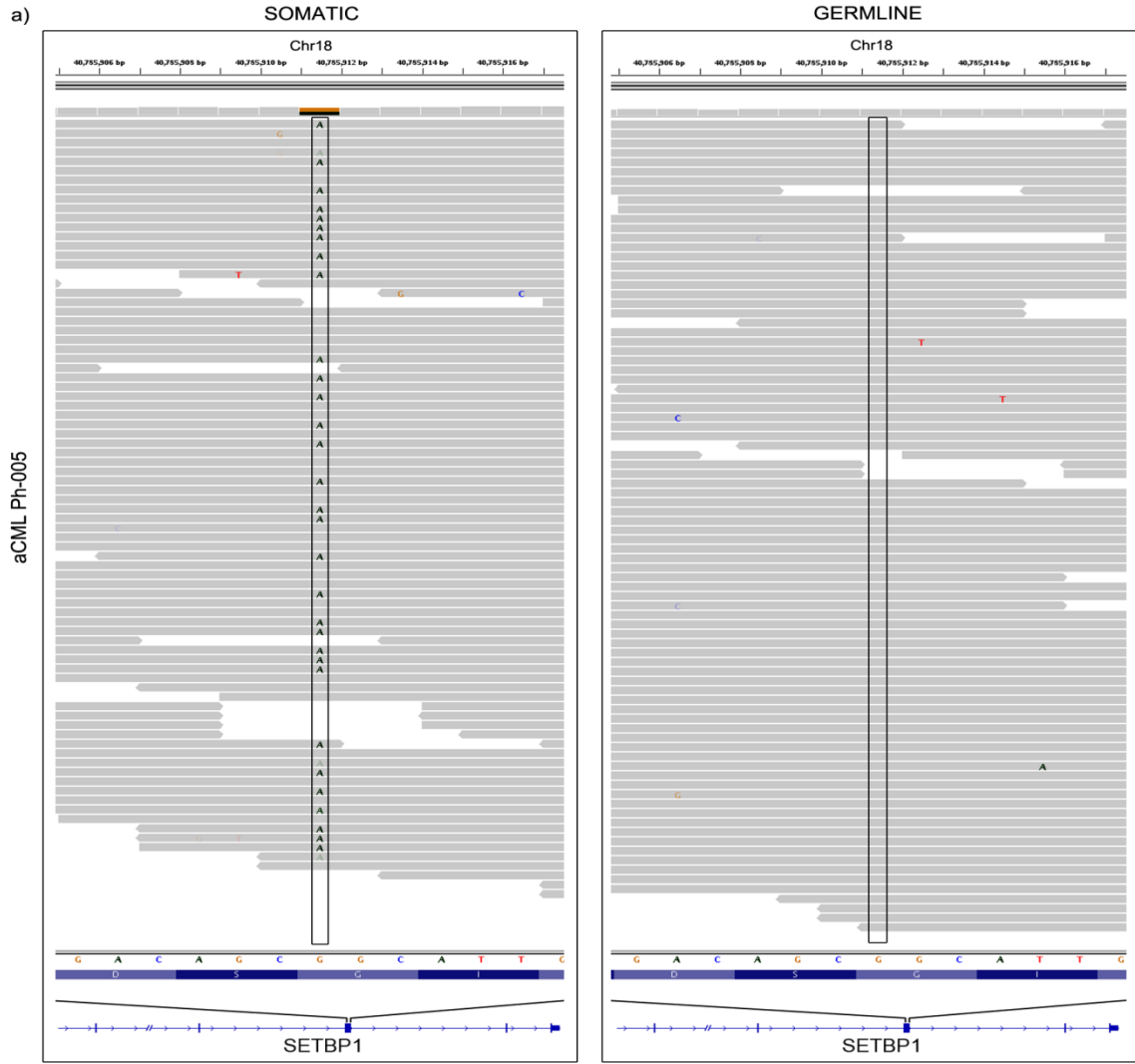
Focusing our attention on recurrently mutated genes, we found *EZH2* variants in 2 patients (subjects 4 and 8) and *SETBP1* alterations in other two (subjects 3 and 5). No other recurrent mutations were identified, even when we decided to consider variations with a frequency <35%. *EZH2* encodes the catalytic subunit of the polycomb repressive complex 2 (PRC2), a

highly conserved histone H3 lysine 27 (H3K27) methyltransferase that influences stem cell renewal by epigenetic repression of genes involved in cell fate decisions (93). *EZH2* mutations were previously identified in different myeloid neoplasias, most commonly in myelodysplastic/myeloproliferative neoplasms and were associated with a poor prognosis in patients showing these alterations, compared to the patients without them (93). The second recurrent mutated gene was *SETBP1*: two of the eight patient samples analyzed showed a somatic substitution causing Gly870Ser mutation, with a frequency of 53% in the first (subject 3) and a frequency of 47% in the second patient (subject 5).

Recurrent SETBP1 mutations

After the identification of recurrent, somatic *SETBP1* variants in aCML patients, we focused our study on the characterization of these mutations, considering that this gene was not previously correlated to cancer and its function was largely unknown. After the validation of this particular mutation using the traditional Sanger sequencing (fig 10), we decided to re-sequence *SETBP1* in a larger panel of samples, including subjects with aCML or other hematological malignancies and in cell lines representative of the most common human solid cancers.

a)



b)

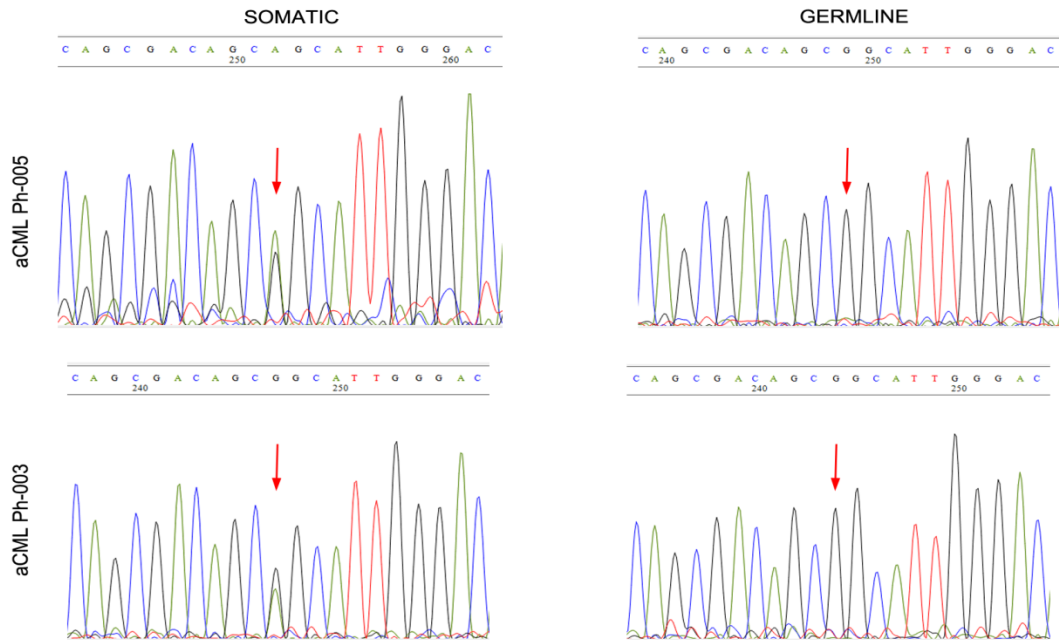


Figure 10. SETBP1 somatic substitution causing Gly870Ser mutation. a) Detail of the whole-exome sequencing reads from patient Ph-005 showing a G>A substitution in 47% of the reads in the leukemic (left panel) sample but not in the germline genomic DNA (right panel). The mutation is shown in black. The intensity of the color is proportional to the corresponding read quality. The black boxes highlight the position of the mutated nucleotide. In the lower panels, the SETBP1 genomic locus, the wild-type nucleotide sequence and the corresponding amino acid sequence are shown. b) Confirmation of the whole-exome sequencing results obtained using Sanger Sequencing in patients Ph-005 and Ph-003. The red arrows point to the mutated nucleotide. (R Piazza et al. Recurrent SETBP1 mutations in atypical chronic myeloid leukemia. Nat Genet 2013;45:18-24)

As reported in table 11, using this approach, we detected mutations of *SETBP1* in 17 of 70 aCML cases, (24.3%), in 3 of 30 unclassified MDS/MPN cases (10%), in 3 of 82 CMML cases (4%), and in 1 of 4 CNL cases (25%) analyzed. No evidence of *SETBP1* mutations could be found in other hematological malignancies samples analyzed, represented by AML, ALL, CLL, MDS, CML, PMF, PV, ET, JMML cases.

Tumor type	Nr.of samples	Nr of mutated samples	% mutated samples
AML	106	0	0
ALL	62	0	0
CLL	32	0	0
MDS	100	0	0
MPN			
CML	42	0	0
PMF	33	0	0
PV	42	0	0
ET	36	0	0
CNL	4	1	25%
MDS/MPN			
aCML	70	17	24%
MDS/MPN-u	30	3	10%
CMML	82	3	4%
JMML	5	0	0
Cell lines			
Brain/CNS	11	0	0
Breast	24	0	0
Colorectal	28	0	0
Esophagus/Stomach	20	0	0
Head & Neck	8	0	0
Liver/Bladder	15	0	0
Lung: NSCLC	103	0	0
Lung: SCLC	17	0	0
Melanoma/Skin	25	0	0

Mesothelioma	15	0	0
NHL	12	0	0
Others (Kidney, Thyroid, Testis)	9	0	0
Ovarian	8	0	0
Pancreas	11	0	0
Prostate	7	0	0
Sarcoma	18	0	0
Uterus	13	0	0
TOTAL	988	24	

Table 11. Frequency of SETBP1 mutations in 644 patient samples and 344 cancer cell lines. Diseases are grouped according to WHO classification (2008). A detailed list of cell lines is reported in Table 10. (R Piazza et al. Recurrent SETBP1 mutations in atypical chronic myeloid leukemia. Nat Genet 2013;45:18-24)

Taken together, considering the presence of *SETBP1* alterations only in aCML, MDS/MPN, CMML and CNL cases and not in 458 individuals with other hematological diseases or in 344 cell lines analyzed (table 11), these results suggested that *SETBP1* mutations can be specifically correlated to aCML and closely related disorders. Constitutive DNA was available from four of seventeen *SETBP1*-mutated aCML cases, the analysis of which showed that mutations encoding p.Glu858Lys, p.Asp868Asn, p.Gly870Ser and p.Ile871Thr alterations were somatically acquired.

In total, we identified 24 SETBP1 mutated patients: of these, 22 (92%) were located in a region comprised between Glu858 and Ile871 (Fig 11), within the SETBP1 SKI homologous region (94).

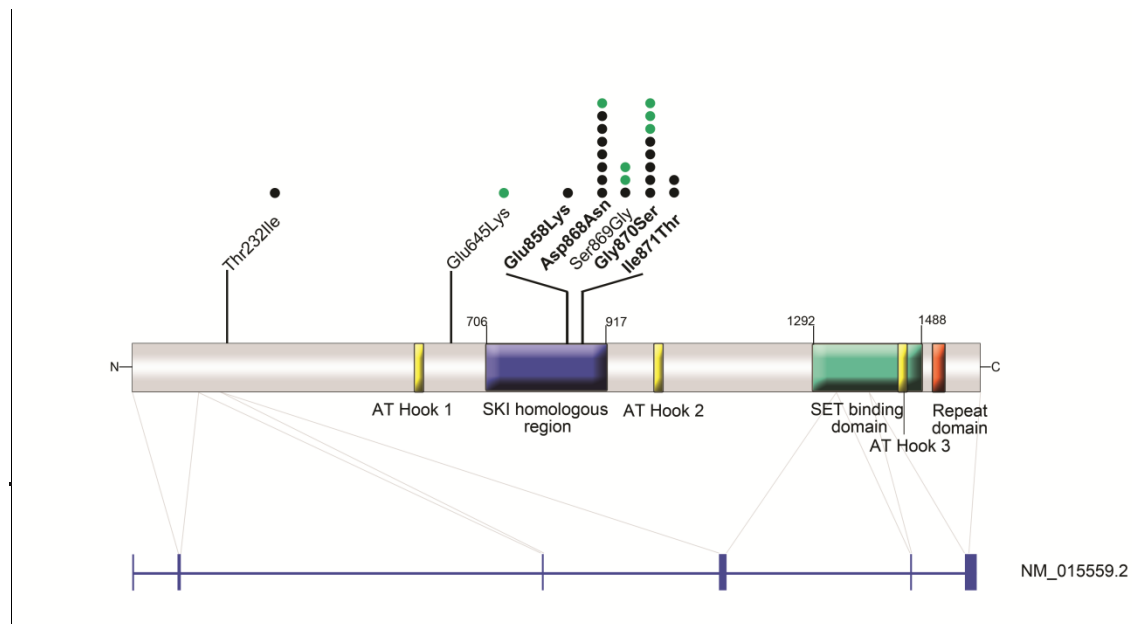


Figure 11. Distribution of alterations on the SETBP1 protein. Five exons (blue bars) encode isoform A of the protein (1,596 amino acids). The SETBP1 sequence contains three AT hook domains (amino acids 584-596, 1,016-1,028, 1,451-1,463), a SKI homologous region (amino acids 706-917), a SET-binding domain (amino acids 1,292-1,488) and a repeat domain (amino acids 1,520-1,543). Altered amino acids identified in our analysis are highlighted: black circles represent alterations found in aCML samples, and green circles represent alterations found in other diseases. Variants confirmed as somatic are indicated in bold. SETBP1 numbering refers to the NCBI reference sequence NM_015559.2. (R Piazza et al. Recurrent SETBP1 mutations in atypical chronic myeloid leukemia. Nat Genet 2013;45:18-24)

As reported in Fig 11, some of these variations were more frequent than the others, specifically p.Glu858Lys, p.Asp868Asn, p.Ser869Gly, p.Gly870Ser and p.Ile871Thr. We analyzed these alterations using two specific software, named SIFT and PolyPhen-2, in order to understand how these mutations could potentially affect the function of the protein. The first software, named “Sorting Tolerant From Intolerant”, SIFT, assesses the effect of a substitution, assuming that important positions in a protein sequence have been conserved throughout evolution and therefore substitutions at these positions may affect protein function. So, by using sequence homology, SIFT predicts the effects of all possible substitutions at each position in the protein sequence (95). PolyPhen-2, instead, calculates the naive Bayes posterior probability that a given mutation is damaging and reports estimates of false positive (the chance that the mutation is classified as damaging when it is in fact non-damaging) and true positive (the chance that the mutation is classified as damaging when it is indeed damaging) rates. A mutation is also appraised qualitatively, as benign, possible damaging or probably damaging (96). The results for all the SETBP1 protein variants invariably predicted a damaging alteration of the normal protein function.

Another analysis that we performed was to assess the relative expression levels of the two alleles by deep sequencing in the cDNA of three aCML cases, obtaining a frequency of the mutated allele of 79%, 45% and 38%, and an absolute coverage for the mutated bases of 905, 440 and 523, respectively. This frequency is compatible with a somatic heterozygous status without substantial imbalance in allelic expression (81). An increasing number of cancer genes have been found to carry recurrent somatic mutations in hematological malignancies, including genes involved in signal transduction (*NPM1*, *JAK2*, *KRAS*, *NRAS*, *CBL*); DNA

methylation (*DNMT3A*, *JARID2*, *TET2*, *IDH1*, *IDH2*); transcriptional regulation (*RUNX1*, *SUZ12*, *EED*, *CEBPA*, *FLT3*); chromatin modification (*EZH2*, *ASXL1*, *RBBP4*) and, most recently, RNA splicing (*SF3B1*) (97). To study the relationship between *SETBP1* mutations and these particular oncogenes involved in myeloid malignancies, we analyzed the possible presence of mutations in these genes in a cohort of 61 aCML patients, including 14 with *SETBP1* mutations and 47 without. As shown in Fig12, no significant association or mutual exclusion with *SETBP1* mutations was observed (81).

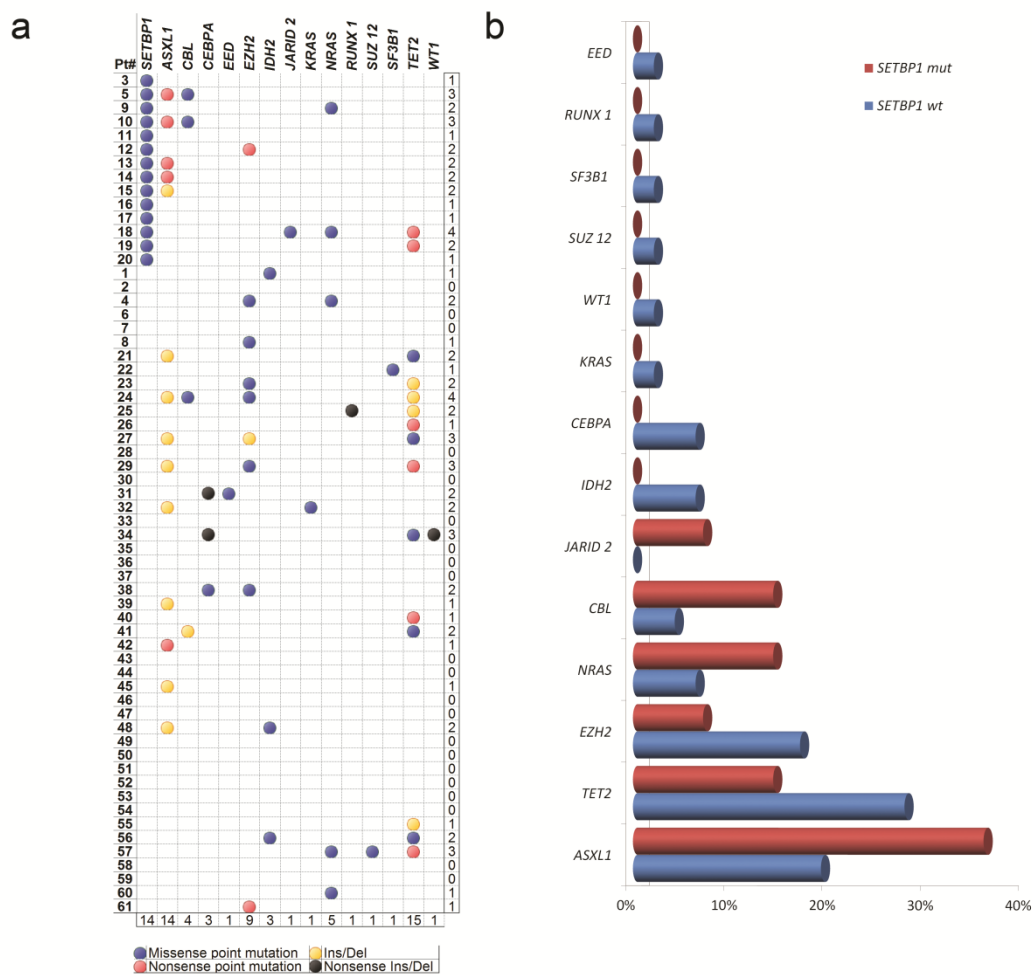


Figure 12. Mutation profile of 61 aCML cases for a panel of 15 genes. a) For each case and each gene are reported total numbers of mutations. Different colors indicate different kinds of mutations. On the right are reported total numbers of mutations. b) Distribution of mutations in aCML. For each gene, the percentage of mutations associated with either wild-type or mutated SETBP1 is reported. (R Piazza et al. Recurrent SETBP1 mutations in atypical chronic myeloid leukemia. Nat Genet 2013;45:18-24)

However, we observed that *ASXL1* mutations were present more frequently in cases mutated, whereas *TET2* mutations in cases with wild-type *SETBP1*. These are, however, preliminary data and it will be important to extend the analysis to a larger panel of patients to conclude if these differences are truly significant.

RNA-seq analysis

The transcriptome is the complete set of transcripts in a cell. Understanding the transcriptome is essential for interpreting the functional elements of the genome and revealing the molecular constituents of cells and tissues, and for understanding development processes and diseases (98). Different technologies have been developed in recent years in order to characterize and quantify the transcriptome, but only recently RNA-Seq (RNA-Sequencing) has revolutionized the manner in which transcriptomes are analyzed.

RNA-Seq is a novel high-throughput DNA sequencing method that uses deep-sequencing technologies, allowing the entire transcriptome to be surveyed in a very high-throughput and quantitative manner (98) and providing the ability to profile alternatively spliced gene transcripts, post-transcriptional changes, gene fusions, mutations and changes in gene expression. In order to identify possible chimeric fusion genes due to cryptic chromosomal rearrangements, we decided to apply RNA-Seq to seven *SETBP1*-mutated aCML cases and six *SETBP1*-wild-type aCML cases. Data generated from this analysis were analyzed using our in-house software named FusionAnalyzer (88). This software is able to identify driver rearrangements from transcriptome paired-end data even in presence of single nucleotide mismatches, such as single nucleotide polymorphisms or sequencing artifacts in the context of the breakpoint region or in presence of extremely low-coverage data, so that we proposed it as a potent and practical tool for the identification of functional rearrangements in the context of high-throughput transcriptome sequencing data (88). According with the literature, showing that aCML is typically not associated to gross genomic alterations, using this method, we didn't identify fusion genes. However, RNA-Seq results were used to validate the expression of the mutated *SETBP1*, confirming that transcript levels of *SETBP1* were similar in mutated and wild-type cases.

Clinical course of aCML cases according to SETBP1 mutation status

We obtained clinical information for 14 *SETBP1*-mutated and 24 *SETBP1*-wild type cases. We analyzed them by univariate analysis, considering specific parameters: sex, age, white blood cell count, hemoglobin concentration and platelet number at diagnosis, the percentage of peripheral blood blasts and overall survival. By using this approach we demonstrated that *SETBP1*-mutated cases were associated with a worse prognosis (median survival=22 versus 77 months, $P=0.01$, hazard ratio=2.27) as indicated in Fig.13a and had a higher number of white blood cells at diagnosis (median of 81.0 versus 38.5×10^9 cells per liter, $P=0.008$ Fig.13b) if compared with *SETBP1*-wild type cases (81). No significant differences could be demonstrated for number of peripheral blood blasts, age, hemoglobin concentration or platelet count, as shown in Fig 13c-d, and no difference was observed in sex distribution between the two groups, too. The negative effect of *SETBP1* mutations on survival was present ($P=0.035$) even when we decided to filter data considering the effects of age and white blood cell count or of sex, hemoglobin concentration, platelet number or blasts count in peripheral blood.

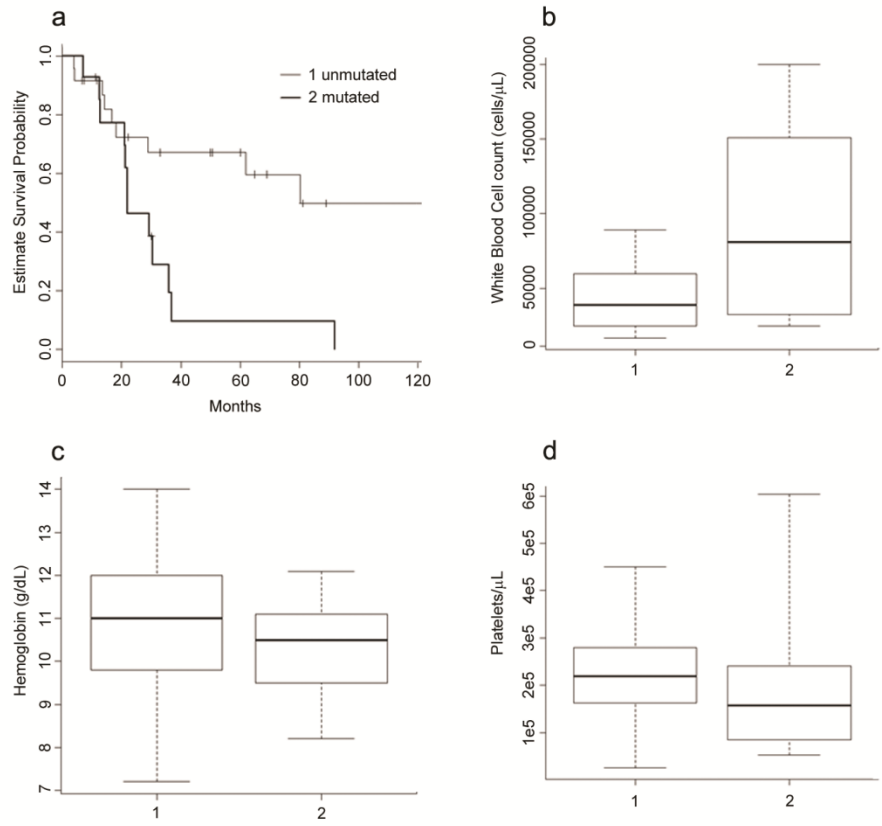


Figure 13. Clinical findings in cases with wild-type and mutated SETBP1 . a) overall survival; b) white blood cell count; c) hemoglobin concentration; d) platelet number in 14 aCML cases with mutated SETBP1 and 24 cases with wild-type SETBP1. Values are shown as median (horizontal line), 25th and 75th percentiles (boxes), and maximum-minimum ranges (dotted lines). Errors bars, s.e.m. (R Piazza et al. Recurrent SETBP1 mutations in atypical chronic myeloid leukemia. Nat Genet 2013;45:18-24)

Biological effects of the SETBP1 p.Gly870Ser alteration

SETBP1 is a poorly characterized protein of 1542 amino acids and 170 KDa, localized predominantly in the nucleus; the physiologic function as well as the molecular mechanism by which SETBP1 acts remain largely unknown. The protein contains a region homologous to the dimerization domain of SKI and a SET binding region (99). It is known that the protein SET (I2PP2A) inhibits PP2A through the phosphorylation of the tyrosine 307 of PP2A. In fact, SET is overexpressed in multiple solid tumors and in CML, where its correlation has been demonstrated with the expression and oncogenic activity of BCR/ABL, leading to PP2A inhibition (100). In addition it has been recently suggested that overexpression of SETBP1 promotes an increase in full-length SET levels, impairing the phosphatase activity of the tumor suppressor PP2A through the formation of a SETBP1-SET-PP2A complex, and promoting the proliferation of cells (99) (Fig 14).

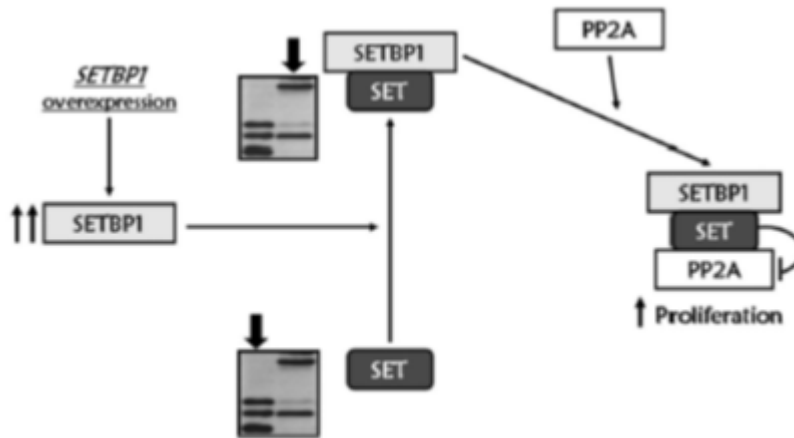


Figure 14. Proposed molecular model for SETBP1 signaling pathway. (I Cristòbal, et al. SETBP1 overexpression is a novel leukemogenic mechanism that predicts adverse outcome in elderly patients with acute myeloid leukemia. *Blood* (2010);115:615-625)

The SETBP1 sequence in which the mutations cluster is highly conserved throughout evolution among vertebrates, which suggests an important yet unknown biological role for this region. We decided to analyze this region using a particular server, named ELM (Eukaryotic Linear Motif) server, that is a new bioinformatics resource for investigating candidate short non-globular functional motifs in eukaryotic proteins, that is implemented with several logical filters to eliminate false positives (101). Considering the prediction of the ELM server, this region represents a virtually perfect degron (a specific sequence of amino acids in a protein that directs the initial step of degradation), containing the consensus binding region (DpSGXXpS/Pt, where pS and pT represent phosphorylated residues) for β -

TrCP1, the substrate recognition subunit of the E3 ubiquitin ligase (amino acids 868-873; fig 15) (81).

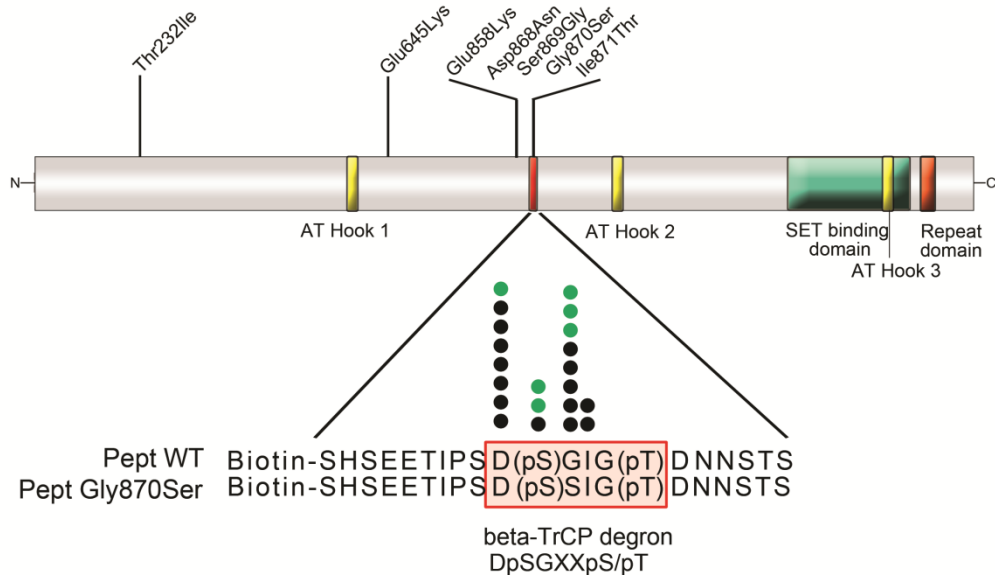


Figure 15. The *b-TrCP1* degron motif (amino acids 868-873) is reported in red on the *SETBP1* protein scheme. The sequences of biotinylated phosphorylated peptides (amino acids 859-879) used in the experiments are given. Black circles indicate alterations found in *aCML*; green circles indicate alterations found in other diseases. (R Piazza et al. Recurrent *SETBP1* mutations in atypical chronic myeloid leukemia. *Nat Genet* 2013;45:18-24)

This degron is part of a PEST domain (amino acids 860-884, HSEETIPSDSGIGTDNNSTSDQAEK), that is typically associated with proteins that have an extremely short intracellular half-life. So, the presence of the degron and the PEST domain suggests that this region might play an

important role for ubiquitin binding and for subsequent protein degradation. To demonstrate this hypothesis, we performed peptide pulldown experiments using biotinylated phosphorylated peptides encompassing the SETBP1 region (amino acids 859-879) of either wild-type or Gly870Ser protein (Innovagen; details reported in Material and Methods). Whereas the wild-type peptide, when incubated with the TF1 cell lysate, could efficiently bind β -TrCP1, the peptide carrying the Gly870Ser mutation was incapable to do it, suggesting a possible difference in SETBP1 protein stability due to this alteration (fig 16).

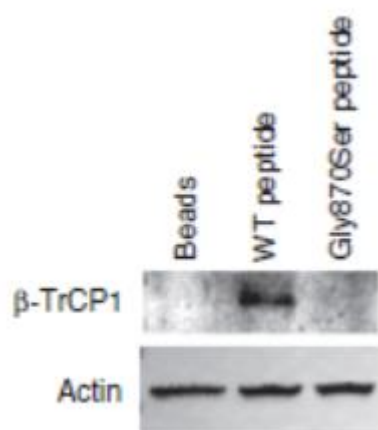


Figure 16. Peptide pulldown experiments performed using TF1 total cell lysate and phosphorylated peptides. Beads with no peptides were used to control for nonspecific binding. Immunoblotting for β -TrCP1 was performed on the bound fractions. Actin was used as a loading control. (R Piazza et al. Recurrent SETBP1 mutations in atypical chronic myeloid leukemia. Nat Genet 2013;45:18-24)

It is known that for the binding of β -TrCP1 the presence of one phosphorylated serine and one phosphorylated threonine in the core consensus region of the degron is necessary. To demonstrate the specificity of the interaction that we observed in the first pulldown experiment, we decided to perform the same experiment using, this time, dephosphorylated peptides and purified recombinant β -TrCP1. In absence of phosphorylation of these two residues, the interaction between the wild-type peptide and β -TrCP1 did not occur (Fig17).

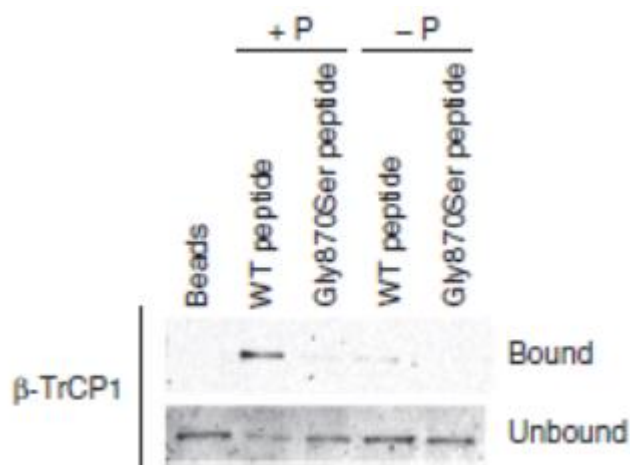


Figure 17. Peptide pulldown experiment performed using recombinant SCF- β -TrCP1 complex on phosphorylated (+P) and dephosphorylated (-P) peptides representing either wild-type SETBP1 or mutated SETBP1. Beads with no peptide were used to control for nonspecific binding. Immunoblotting for β -TrCP1 was performed on bound as well as unbound (control) fractions. (R Piazza et al. Recurrent SETBP1 mutations in atypical chronic myeloid leukemia. Nat Genet 2013;45:18-24)

These data suggest that the presence of Gly870Ser alteration deleted a site critical for ubiquitination preventing the subsequent degradation of the SETBP1 protein, leading to an

increase in protein stability. To test this hypothesis, we assessed the expression and stability of SETBP1 protein using a specific antibody in TF1 cells transduced with viruses expressing either wild-type or SETBP1 Gly870Ser and expressing comparable levels of *SETBP1* mRNA. As shown in Fig.18a, as we expected, SETBP1 protein was nearly undetectable in cells expressing wild-type SETBP1, in line with its supposed very short half-life, whereas cells expressing SETBP1 Gly870Ser showed higher levels of SETBP1 protein, recognized as a band of approximately 250 kDa, co-migrating with the positive control represented by whole normal fetal stomach lysate.

It has been reported that *SETBP1* overexpression protects SET from protease cleavage, increasing the amount of full-length SET protein, and leading to the formation of a SETBP1-SET-PP2A complex that results in PP2A inhibition, promoting proliferation of the leukemic cells (99). In order to assess the effect of SETBP1 mutation on SET protein and PP2A activity, we performed immunoblotting experiments for detection of SET, phosphorylated PP2A (pPP2A) and PP2A in TF1 cells expressing wild-type SETBP1 or SETBP1 Gly870Ser. We observed a greater amount of SET protein in SETBP1-mutated TF1 cells (fig 18b), despite the presence of similar levels of mRNA between wild-type and mutated SETBP1 TF1 cell lines, and significantly reduced PP2A activity in the cells expressing the mutation (fig 18c), as well as greater PP2A phosphorylation at position Tyr307, a well-known marker of PP2A inactivation (fig 18b) (81). We investigated also the effect of the SETBP1 mutation on the proliferation rate comparing wild-type and mutated SETBP1 expressing TF1 cells; as reported in fig 18d, showing that SETBP1 Gly870Ser cells had a higher proliferation rate respect to

SETBP1 wild-type cells, when cultured adding the standard concentration of granulocyte-macrophage colony-stimulating factor (GM-CSF) to the medium.

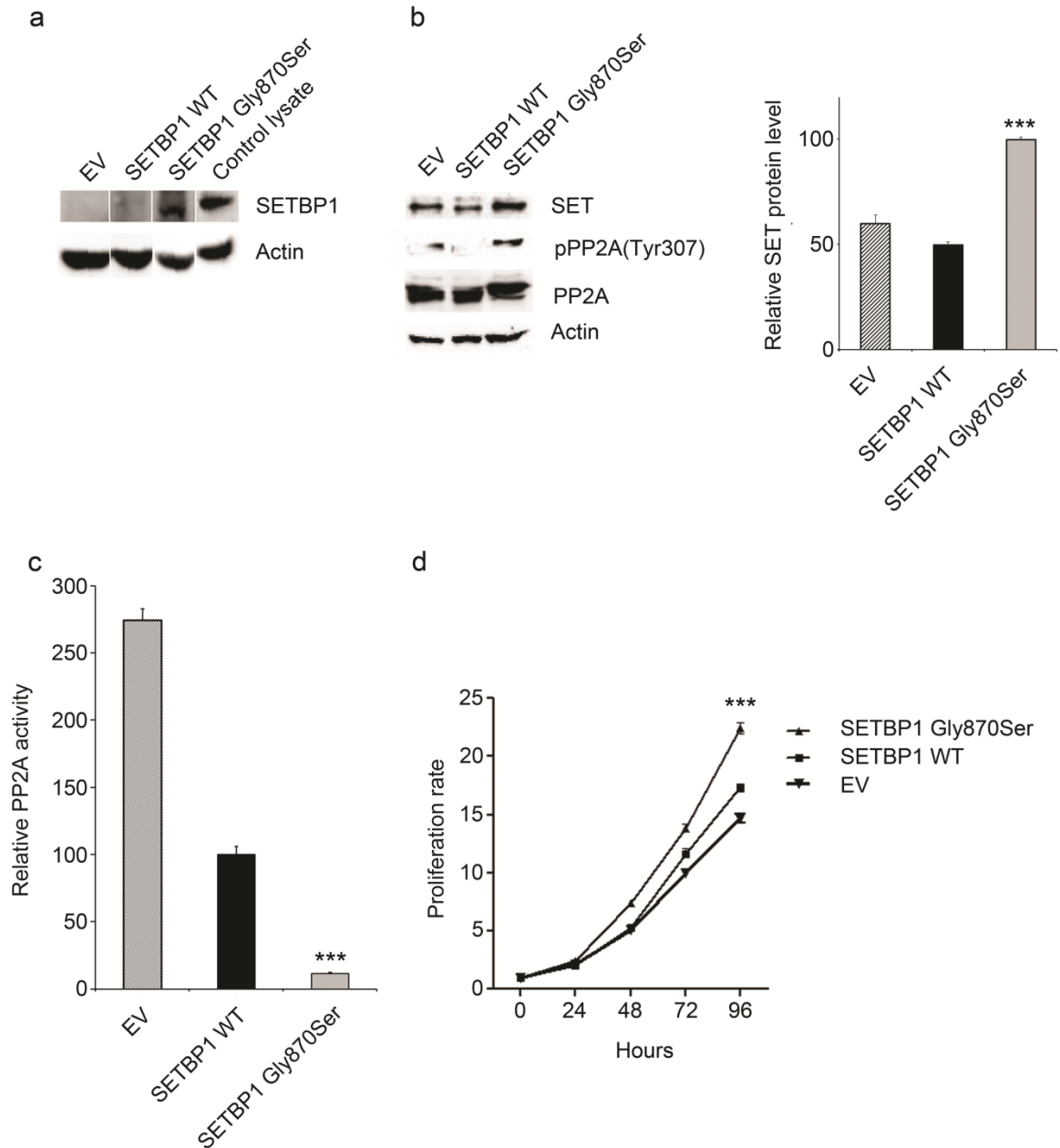


Figure 18. Effects of the SETBP1 Gly870Ser alteration on SETBP1 and SET protein expression, PP2A activity and cell growth. TF1 cells transfected with empty vector (Pmigr1, ev) OR VECTOR expressing wild-type SETBP1 or SETBP1 Gly870Ser. a) Immunoblotting for SETBP1 on whole cell lysate. Whole normal fetal stomach lysate was used as a control for SETBP1 protein; actin was used as loading control. b) immunoblotting for SET, phosphorylated PP2A (Ppp2a), PP2A and actin on whole-cell lysates. c) Lysates were used to assess the activity of PP2A. d) Growth rate of the cells as measured by tritiated thymidine incorporation. (R Piazza et al. Recurrent SETBP1 mutations in atypical chronic myeloid leukemia. Nat Genet 2013;45:18-24)

Intracellular localization of SETBP1

We know that *SETBP1* encodes a protein of 170 kDa, localized predominantly in the nucleus (99). To investigate possible differences in the intracellular localization of SETBP1, wild-type protein and SETBP1 Gly870Ser forms fused with GFP were introduced into the 293T human cell line and sorted to express similar levels of protein (fig 19) (81).

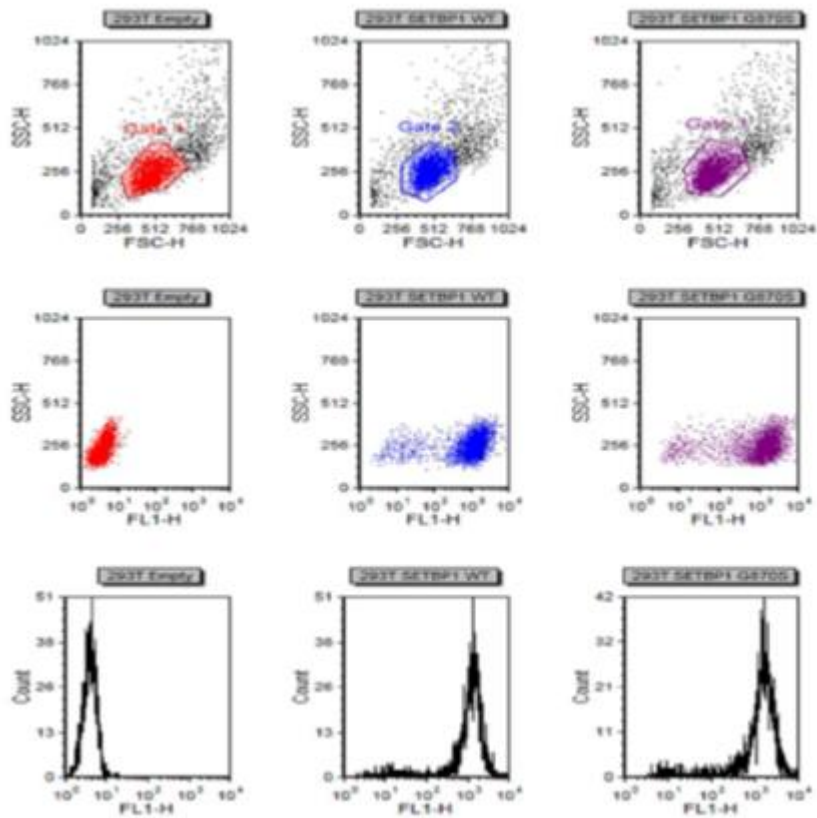


Figure 19. Flow cytometry analysis of the 293T cells expressing the GFP-fused SETBP1 wild-type protein (middle column), the Gly870Ser variant (right column) or a negative control (left column). (Piazza et al. Recurrent SETBP1 mutations in atypical chronic myeloid leukemia. *Nat Genet* 2013;45:18-24)

Using confocal microscopy, it was possible to analyze the intracellular localization of SETBP1, in its wild-type and mutated forms. As shown in Fig. 20, the mutated form of SETBP1 maintained the normal nuclear localization, giving no significant alteration in the intracellular presence of the protein. The only difference that was possible to observe was a more

punctate appearance showed by cells expressing SETBP1 mutated form, respect to the wild-type expressing cells. However, further analyses are ongoing, to determine if there are other differences, focusing the attention also on SET and PP2A localization.

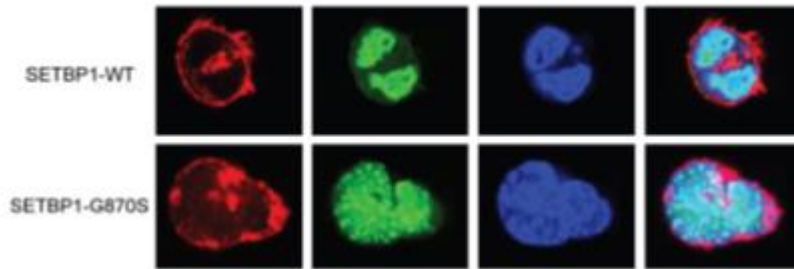


Figure 20. Intracellular localization of wild-type and Gly870Ser SETBP1 in 293T cells transfected with a GFP-SETBP1 (wt or mutated) construct. The four columns show Falloidin (red), GFP (green), TOTO-3 iodide (blue) fluorescence and the product of the combined channels, respectively. (R Piazza et al. Recurrent SETBP1 mutations in atypical chronic myeloid leukemia. Nat Genet 2013;45:18-24)

Discussion

Atypical Chronic Myeloid Leukemia (aCML) is a heterogeneous disorder belonging to the group of myelodysplastic/myeloproliferative (MDS/MPN) syndromes. This disease shares clinical features with CML, as splenomegaly and myeloid predominance in the bone marrow, with some dysplastic features but without a differentiation block. From a laboratory point of view aCML shows the same abnormalities, such as myeloid proliferation and low leukocyte alkaline phosphatase values. However, lack of the pathognomonic Philadelphia chromosome and of the resulting BCR-ABL fusion point to a different pathogenetic process. Because no specific recurrent genomic or karyotypic abnormalities have been identified in aCML, the molecular pathogenesis of this disease has remained elusive and the outcome dismal (median survival of 37 months after diagnosis) (5), with no improvement over the last 20 years. This prognosis sharply contrasts with the outcome for CML, for which the prognosis was markedly improved by the development of imatinib a specific inhibitor of BCR-ABL protein (102).

Considering the lack of information about aCML, and because no specific recurrent genomic or karyotypic abnormalities have been correlated until now to this disease, *SETBP1* represents the first gene shown to be enriched and recurrently mutated in aCML (81). The identification of *SETBP1* mutations in approximately one-quarter of aCML cases analyzed using next-generation sequencing suggests an important role of this gene in the pathogenesis of this disease. In addition, the finding of *SETBP1* alterations in CMML and CNL cases, and especially in the closely related MDS/MPN disorders suggests that these mutations may play a specific role in the pathogenesis of aCML and related disorders, such as MDS/MPN. Furthermore, considering that cases with *SETBP1* mutations had a worse prognosis compared

with cases with wild-type *SETBP1*, this gene can be considered the first prognostic tool available for the stratification of aCML patients. Notably, in a recent work (103) whole exome sequencing has been performed in a patient with myelodysplastic syndrome before and after progression to acute myeloid leukemia, leading to the identification of *SETBP1* mutations following leukemic transformation, which suggests that mutations on this gene may play also a significant role in MDS progression.

Little is known about *SETBP1*, except that it is located at chromosome 18q21.1 and encodes a protein of 1,596 residues (NM_015559.2, long isoform) initially identified as a 170 kDa nuclear protein that binds to *SET* (94) and that its activating insertions support the recovery of granulopoiesis in chronic granulomatous disease (104). *SETBP1* physiological function as well as the molecular mechanism by which *SETBP1* acts remain unclear. It has been reported that *SETBP1* overexpression is a recurrent molecular event with poor prognostic impact in AML (99) and that it protects *SET* from protease cleavage, increasing the amount of full-length *SET* protein, and leading to the formation of a *SETBP1*-*SET*-*PP2A* complex (99). In particular, the formation of this complex causes the inhibition of the *PP2A* phosphatase promoting the proliferation and expansion of leukemic cells (99). Our experimental data confirmed the nuclear localization of *SETBP1*, but indicated a greater molecular weight respect to the predicted of 170 kDa. Up to now, the reason for this is not clear, so extensive additional work will be necessary to clarify this point. Our idea is, however, that this change can be due to possible post-translational modifications.

SETBP1 has been reported as a chromosomal translocation partner in rare cases of lymphoid and myeloid acute leukemias (105), and most recently, it has been suggested that *SETBP1*

mutations are associated with specific cytogenetic aberrations involving chromosomes 3 and 7 in MDS and secondary AML (sAML) (106).

CMML is a nosological entity considered very similar to aCML with largely overlapping diagnostic criteria. The finding that *SETBP1* somatic mutations are present in cases diagnosed as CMML with a prevalence almost seven times lower than in aCML represents the first biological difference that can be used to distinguish these two diseases and suggests that in the next WHO revision, new criteria taking into account the presence of *SETBP1* mutations must be included in order to properly discriminate between these two disorders.

In myelodysplastic syndromes an increasing number of cancer genes have been found to carry recurrent somatic mutations, including genes involved in signal transduction (*KRAS*, *NRAS*), DNA methylation (*IDH1/2*), transcriptional regulation (*CEBPA*, *SUZ12*), chromatin modification (*ASXL1*) and RNA splicing (*RBBP4*), and it may be important to study a possible correlation between all these mutations and the progression of this group of diseases. Also in aCML, mutations of known oncogenes such as *NRAS*, *KRAS*, *TET2*, *EZH2* and *CBL* have been described (78); our experiments, performed in order to test the presence of a potential association between *SETBP1* variants and a large number of these mutations, suggest that there is no statistically significant association or mutual exclusion. Considering the cases analyzed, the only results obtained are a trend towards a positive association of *ASXL1* and a negative association of *TET2* variants with the mutated *SETBP1* phenotype. Further analyses with larger cohorts will be necessary to determine if these differences are real.

The only known interactions of SETBP1 are with the *HOXA9* and *HOXA10* promoters (107) and with SET through its SET-binding domain (94). It has been demonstrated that Setbp1 overexpression can efficiently immortalize myeloid progenitors in culture, and is also essential to sustain the self-renewal of these immortalized cells, identifying it as a novel regulator of LSC (Leukemic Stem Cells) self-renewal (107). In particular, there is the evidence that SETBP1 could function as a transcription factor, contributing to leukemia transformation via the activation of *HOXA9* and *HOXA10*. It has been proposed, in addition, that SETBP1 expression may contribute to myeloid leukemia development by inhibition of PP2A activity, forming a complex with SET and PP2A, enhancing the stability of SET and its inhibition of PP2A (99). For this reason we tested whether the expression of SETBP1 Gly870Ser could result in SET stabilization and PP2A inhibition, showing that the presence of the mutated SETBP1 caused accumulation of SET protein and phosphorylation of the PP2A phosphatase at position Y307, a well-known marker of PP2A inactivation.

The SETBP1 region where the somatic alterations cluster is highly conserved among vertebrates, suggesting its possible crucial biological role. In particular, using the ELM server, we discovered that this region represents a virtually perfect degron (a specific sequence of amino acids in a protein that directs the initial step of degradation), containing the consensus binding region for β -TrCP1, the substrate recognition subunit of the E3 ubiquitin ligase (amino acids 868-873). This region, in other words, appears important for ubiquitin binding and the subsequent SETBP1 degradation. So, the dysregulation of SETBP1 protein levels and activity that we observed can be explained by the functional loss of this degron binding region for the E3-Ubiquitin Ligase β -TrCP in the mutant form of SETBP1 protein, causing the

escape of SETBP1 from proteasomal degradation and accumulation of the mutated protein. Although we tested only SETBP1 Gly870Ser, it is reasonable to speculate that a virtually identical mechanism of action is shared across all the variants occurring inside the degnon conserved region (81).

RNA-seq analysis performed in aCML cells expressing wild-type and mutated *SETBP1* showed that several TGF- β target genes were differentially expressed. This result is in line with the known activity of SKI (and possibly of the SKI homology domain of SETBP1 protein) on TGF- β via its interaction with SMADs (108). Further studies will be required to confirm this hypothesis.

Several germline *SETBP1* mutations have been reported previously in a de novo germline syndrome, named Schinzel-Giedion syndrome (SGS): SGS is a rare multiple congenital malformation syndrome defined by characteristic facial features, profound developmental delay, severe growth failure, and multiple congenital anomalies (109). In 2010, by exome sequencing, *SETBP1* was identified as the gene responsible for SGS. The mutations identified in *SETBP1* were all heterozygous variants occurring in the *SKI* homologous region and all the affected residues were highly conserved throughout the evolution. Strikingly, the mutations identified in our study are virtually superimposable to those of SGS. There are two indications that the SETBP1 alterations seen in SGS and aCML may result in a gain-of-function effect (110). First, a gain-of-function mechanism has been reported for other syndromes, for example Noonan syndrome (111), that have very similar mutations clustering in *PTPN11* and *SOS1*. Second, the phenotype of individuals with partial chromosome 18q deletions which affect *SETBP1* does not resemble Schinzel-Giedion syndrome. In addition *SETBP1* shows

ubiquitous expression (112), which is consistent with the multi-systemic defects in patients affected by the Schinzel-Giedion syndrome. It is tempting to connect the SGS phenotype, and so the pathogenesis of aCML showing the same mutations, to alterations in the TGF- β pathway, knowing the important role of this cytokine in bone formation and remodeling (113), but additional research is essential to confirm this idea. It is known that SGS is a very debilitating syndrome, and in most cases patients die in the perinatal period; when they survive, some develop tumors, predominantly of neuroepithelial origin (114). Until now there is no evidence of a particular predisposition to myeloid malignancies in SGS cases, but the number of cases reported is small so we can't exclude this possibility.

In summary, using -exome sequencing we identified *SETBP1* mutations in approximately one-quarter of aCML cases and these patients showed a worse clinical course. This is the first time that recurrent *SETBP1* mutations are correlated to cancer. We demonstrated, in particular, that TF1 cells expressing *SETBP1* Gly870Ser alteration showed higher levels of *SETBP1* protein, SET protein stabilization associated with PP2A inhibition and higher proliferation rates. Considering that *SETBP1* represents the first gene shown to be recurrently mutated in aCML, it may constitute a valuable diagnostic and prognostic tool in the differential diagnosis of MDS/MPN syndromes.

Bibliography

1. Bennett JM, Catovsky D, Daniel MT, et al. The chronic myeloid leukaemias: guidelines for distinguishing chronic granulocytic, atypical chronic myeloid, and chronic myelomonocytic leukaemia. Proposals by the French-American-British Cooperative Leukaemia Group. *Br J Haematol.* 1994 Aug;87(4):746-54.
2. Fend F, Horn T, Koch I, et al. Atypical chronic myeloid leukemia as defined in the WHO classification is a JAK2 V617F negative neoplasm. *Leuk Res.* 2008 Dec;32(12):1931-5.
3. Vardiman JW. The World Health Organization (WHO) classification of tumors of the hematopoietic and lymphoid tissues: an overview with emphasis on the myeloid neoplasms. *Chem Biol Interact.* 2010 Mar 19;184(1-2):16-20.
4. Vardiman JW, Bain B, Inbert M, editors. *Atypical chronic myeloid leukemia.* 2001.
5. Kurzrock R, Bueso-Ramos CE, Kantarjian H, et al. BCR rearrangement-negative chronic myelogenous leukemia revisited. *J Clin Oncol.* 2001 Jun 1;19(11):2915-26.
6. Onida F, Ball G, Kantarjian HM, et al. Characteristics and outcome of patients with Philadelphia chromosome negative, bcr/abl negative chronic myelogenous leukemia. *Cancer.* 2002 Oct 15;95(8):1673-84.
7. Sabattini E, Bacci F, Sagrmoso C, et al. WHO classification of tumours of haematopoietic and lymphoid tissues in 2008: an overview. *Pathologica.* 2010 Jun;102(3):83-7.
8. Vardiman JW, Harris NL, Brunning RD. The World Health Organization (WHO) classification of the myeloid neoplasms. *Blood.* 2002 Oct 1;100(7):2292-302.
9. Lorsbach RB, Shay-Seymore D, Moore J, et al. Clinicopathologic analysis of follicular lymphoma occurring in children. *Blood.* 2002 Mar 15;99(6):1959-64.

10. Vardiman JW, Thiele J, Arber DA, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood*. 2009 Jul 30;114(5):937-51.
11. Zerbini M, Soares F. World Health Organization Classification of tumors of hematopoietic and lymphoid tissues, 4th edition, 2008- major changes from the third edition, 2011. *rev assoc med bras* 2011;57(1):66-73.
12. Tefferi A, Lasho TL, Gilliland G. JAK2 mutations in myeloproliferative disorders. *N Engl J Med*. 2005 Sep 29;353(13):1416-7; author reply -7.
13. Fialkow PJ, Jacobson RJ, Papayannopoulou T. Chronic myelocytic leukemia: clonal origin in a stem cell common to the granulocyte, erythrocyte, platelet and monocyte/macrophage. *Am J Med*. 1977 Jul;63(1):125-30.
14. Garcia-Manero G, Faderl S, O'Brien S, et al. Chronic myelogenous leukemia: a review and update of therapeutic strategies. *Cancer*. 2003 Aug 1;98(3):437-57.
15. Kantarjian HM, Keating MJ, Walters RS, et al. Clinical and prognostic features of Philadelphia chromosome-negative chronic myelogenous leukemia. *Cancer*. 1986 Nov 1;58(9):2023-30.
16. Kurzrock R, Kantarjian HM, Shtalrid M, et al. Philadelphia chromosome-negative chronic myelogenous leukemia without breakpoint cluster region rearrangement: a chronic myeloid leukemia with a distinct clinical course. *Blood*. 1990 Jan 15;75(2):445-52.
17. Faderl S, Talpaz M, Estrov Z, et al. The biology of chronic myeloid leukemia. *N Engl J Med*. 1999 Jul 15;341(3):164-72.

18. Deininger MW, Goldman JM, Melo JV. The molecular biology of chronic myeloid leukemia. *Blood*. 2000 Nov 15;96(10):3343-56.
19. Kipreos ET, Wang JY. Differential phosphorylation of c-Abl in cell cycle determined by cdc2 kinase and phosphatase activity. *Science*. 1990 Apr 13;248(4952):217-20.
20. Yuan ZM, Shioya H, Ishiko T, et al. p73 is regulated by tyrosine kinase c-Abl in the apoptotic response to DNA damage. *Nature*. 1999 Jun 24;399(6738):814-7.
21. Lewis JM, Schwartz MA. Integrins regulate the association and phosphorylation of paxillin by c-Abl. *J Biol Chem*. 1998 Jun 5;273(23):14225-30.
22. Pendergast AM, Quilliam LA, Cripe LD, et al. BCR-ABL-induced oncogenesis is mediated by direct interaction with the SH2 domain of the GRB-2 adaptor protein. *Cell*. 1993 Oct 8;75(1):175-85.
23. Diekmann D, Nobes CD, Burbelo PD, et al. Rac GTPase interacts with GAPs and target proteins through multiple effector sites. *EMBO J*. 1995 Nov 1;14(21):5297-305.
24. Voncken JW, van Schaick H, Kaartinen V, et al. Increased neutrophil respiratory burst in bcr-null mutants. *Cell*. 1995 Mar 10;80(5):719-28.
25. Anastasi J, Moinuddin R, Daugherty C. The juxtaposition of ABL with BCR and risk for fusion may come at the time of BCR replication in late S-phase. *Blood*. 1999 Aug 1;94(3):1137-8.
26. Martinelli G, Terragna C, Amabile M, et al. Alu and transposon recognition site sequences flanking translocation sites in a novel type of chimeric bcr-abl transcript suggest a possible general mechanism for bcr-abl breakpoints. *Haematologica*. 2000 Jan;85(1):40-6.

27. Corso A, Lazzarino M, Morra E, et al. Chronic myelogenous leukemia and exposure to ionizing radiation--a retrospective study of 443 patients. *Ann Hematol.* 1995 Feb;70(2):79-82.
28. Bose S, Deininger M, Gora-Tybor J, et al. The presence of typical and atypical BCR-ABL fusion genes in leukocytes of normal individuals: biologic significance and implications for the assessment of minimal residual disease. *Blood.* 1998 Nov 1;92(9):3362-7.
29. Pane F, Frigeri F, Sindona M, et al. Neutrophilic-chronic myeloid leukemia: a distinct disease with a specific molecular marker (BCR/ABL with C3/A2 junction). *Blood.* 1996 Oct 1;88(7):2410-4.
30. Gordon MY, Dowding CR, Riley GP, et al. Altered adhesive interactions with marrow stroma of haematopoietic progenitor cells in chronic myeloid leukaemia. *Nature.* 1987 Jul 23-29;328(6128):342-4.
31. Puil L, Liu J, Gish G, et al. Bcr-Abl oncoproteins bind directly to activators of the Ras signalling pathway. *EMBO J.* 1994 Feb 15;13(4):764-73.
32. Bedi A, Zehnbaauer BA, Barber JP, et al. Inhibition of apoptosis by BCR-ABL in chronic myeloid leukemia. *Blood.* 1994 Apr 15;83(8):2038-44.
33. Zhao R, Tarone G. Presence of the adhesion inhibitory b-1B integrin isoform on CML but not normal progenitors is at least in part responsible for the decreased CML progenitor adhesion. *Blood.* 1997.
34. Marais R, Light Y, Paterson HF, et al. Ras recruits Raf-1 to the plasma membrane for activation by tyrosine phosphorylation. *EMBO J.* 1995 Jul 3;14(13):3136-45.

35. Raitano AB, Halpern JR, Hambuch TM, et al. The Bcr-Abl leukemia oncogene activates Jun kinase and requires Jun for transformation. *Proc Natl Acad Sci U S A*. 1995 Dec 5;92(25):11746-50.
36. Ilaria RL, Jr., Van Etten RA. P210 and P190(BCR/ABL) induce the tyrosine phosphorylation and DNA binding activity of multiple specific STAT family members. *J Biol Chem*. 1996 Dec 6;271(49):31704-10.
37. Chai SK, Nichols GL, Rothman P. Constitutive activation of JAKs and STATs in BCR-Abl-expressing cell lines and peripheral blood cells derived from leukemic patients. *J Immunol*. 1997 Nov 15;159(10):4720-8.
38. Kabarowski JH, Allen PB, Wiedemann LM. A temperature sensitive p210 BCR-ABL mutant defines the primary consequences of BCR-ABL tyrosine kinase expression in growth factor dependent cells. *EMBO J*. 1994 Dec 15;13(24):5887-95.
39. Skorski T, Kanakaraj P, Nieborowska-Skorska M, et al. Phosphatidylinositol-3 kinase activity is regulated by BCR/ABL and is required for the growth of Philadelphia chromosome-positive cells. *Blood*. 1995 Jul 15;86(2):726-36.
40. Gesbert F, Sellers WR, Signoretti S, et al. BCR/ABL regulates expression of the cyclin-dependent kinase inhibitor p27Kip1 through the phosphatidylinositol 3-Kinase/AKT pathway. *J Biol Chem*. 2000 Dec 15;275(50):39223-30.
41. Zou X, Rudchenko S, Wong K, et al. Induction of c-myc transcription by the v-Abl tyrosine kinase requires Ras, Raf1, and cyclin-dependent kinases. *Genes Dev*. 1997 Mar 1;11(5):654-62.

42. Daley GQ, Baltimore D. Transformation of an interleukin 3-dependent hematopoietic cell line by the chronic myelogenous leukemia-specific P210bcr/abl protein. *Proc Natl Acad Sci U S A*. 1988 Dec;85(23):9312-6.
43. Sirard C, Laneuville P, Dick JE. Expression of bcr-abl abrogates factor-dependent growth of human hematopoietic M07E cells by an autocrine mechanism. *Blood*. 1994 Mar 15;83(6):1575-85.
44. Keeshan K, Cotter TG, McKenna SL. High Bcr-Abl expression prevents the translocation of Bax and Bad to the mitochondrion. *Leukemia*. 2002 Sep;16(9):1725-34.
45. Amarante-Mendes GP, Naekyung Kim C, Liu L, et al. Bcr-Abl exerts its antiapoptotic effect against diverse apoptotic stimuli through blockage of mitochondrial release of cytochrome C and activation of caspase-3. *Blood*. 1998 Mar 1;91(5):1700-5.
46. Skorski T, Bellacosa A, Nieborowska-Skorska M, et al. Transformation of hematopoietic cells by BCR/ABL requires activation of a PI-3k/Akt-dependent pathway. *EMBO J*. 1997 Oct 15;16(20):6151-61.
47. Majewski M, Nieborowska-Skorska M, Salomoni P, et al. Activation of mitochondrial Raf-1 is involved in the antiapoptotic effects of Akt. *Cancer Res*. 1999 Jun 15;59(12):2815-9.
48. Evan GI, Brown L, Whyte M, et al. Apoptosis and the cell cycle. *Curr Opin Cell Biol*. 1995 Dec;7(6):825-34.
49. Kim JH, Chu SC, Gramlich JL, et al. Activation of the PI3K/mTOR pathway by BCR-ABL contributes to increased production of reactive oxygen species. *Blood*. 2005 Feb 15;105(4):1717-23.

50. Burke BA, Carroll M. BCR-ABL: a multi-faceted promoter of DNA mutation in chronic myelogenous leukemia. *Leukemia*. 2010 Jun;24(6):1105-12.
51. Kantarjian HM, Smith TL, O'Brien S, et al. Prolonged survival in chronic myelogenous leukemia after cytogenetic response to interferon-alpha therapy. The Leukemia Service. *Ann Intern Med*. 1995 Feb 15;122(4):254-61.
52. Speck B, Cornu P, Jeannet M, et al. [Allogenic bone marrow transplantation in severe aplastic anemia and refractory acute leukemia. 35 cases]. *Schweiz Med Wochenschr*. 1977 Oct 15;107(41):1451-2.
53. Gambacorti-Passerini C. Part I: Milestones in personalised medicine--imatinib. *Lancet Oncol*. 2008 Jun;9(6):600.
54. Druker BJ, Tamura S, Buchdunger E, et al. Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr-Abl positive cells. *Nat Med*. 1996 May;2(5):561-6.
55. Gambacorti-Passerini C, le Coutre P, Mologni L, et al. Inhibition of the ABL kinase activity blocks the proliferation of BCR/ABL+ leukemic cells and induces apoptosis. *Blood Cells Mol Dis*. 1997 Dec;23(3):380-94.
56. Deininger MW, Goldman JM, Lydon N, et al. The tyrosine kinase inhibitor CGP57148B selectively inhibits the growth of BCR-ABL-positive cells. *Blood*. 1997 Nov 1;90(9):3691-8.
57. le Coutre P, Mologni L, Cleris L, et al. In vivo eradication of human BCR/ABL-positive leukemia cells with an ABL kinase inhibitor. *J Natl Cancer Inst*. 1999 Jan 20;91(2):163-8.
58. Kantarjian H, Sawyers C, Hochhaus A, et al. Hematologic and cytogenetic responses to imatinib mesylate in chronic myelogenous leukemia. *N Engl J Med*. 2002 Feb 28;346(9):645-52.

59. O'Brien SG, Guilhot F, Larson RA, et al. Imatinib compared with interferon and low-dose cytarabine for newly diagnosed chronic-phase chronic myeloid leukemia. *N Engl J Med*. 2003 Mar 13;348(11):994-1004.
60. Graham SM, Jorgensen HG, Allan E, et al. Primitive, quiescent, Philadelphia-positive stem cells from patients with chronic myeloid leukemia are insensitive to STI571 in vitro. *Blood*. 2002 Jan 1;99(1):319-25.
61. Talpaz M, Silver RT, Druker BJ, et al. Imatinib induces durable hematologic and cytogenetic responses in patients with accelerated phase chronic myeloid leukemia: results of a phase 2 study. *Blood*. 2002 Mar 15;99(6):1928-37.
62. Sawyers CL, Hochhaus A, Feldman E, et al. Imatinib induces hematologic and cytogenetic responses in patients with chronic myelogenous leukemia in myeloid blast crisis: results of a phase II study. *Blood*. 2002 May 15;99(10):3530-9.
63. Gambacorti-Passerini CB, Gunby RH, Piazza R, et al. Molecular mechanisms of resistance to imatinib in Philadelphia-chromosome-positive leukaemias. *Lancet Oncol*. 2003 Feb;4(2):75-85.
64. Shah NP, Nicoll JM, Nagar B, et al. Multiple BCR-ABL kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast crisis chronic myeloid leukemia. *Cancer Cell*. 2002 Aug;2(2):117-25.
65. Gorre ME, Mohammed M, Ellwood K, et al. Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. *Science*. 2001 Aug 3;293(5531):876-80.

66. Azam M, Seeliger MA, Gray NS, et al. Activation of tyrosine kinases by mutation of the gatekeeper threonine. *Nat Struct Mol Biol.* 2008 Oct;15(10):1109-18.
67. Mauro MJ. Defining and managing imatinib resistance. *Hematology Am Soc Hematol Educ Program.* 2006:219-25.
68. Kantarjian H, Jabbour E, Grimley J, et al. Dasatinib. *Nat Rev Drug Discov.* 2006 Sep;5(9):717-8.
69. Fava C, Kantarjian H, Cortes J, et al. Development and targeted use of nilotinib in chronic myeloid leukemia. *Drug Des Devel Ther.* 2009;2:233-43.
70. Puttini M, Coluccia AM, Boschelli F, et al. In vitro and in vivo activity of SKI-606, a novel Src-Abl inhibitor, against imatinib-resistant Bcr-Abl+ neoplastic cells. *Cancer Res.* 2006 Dec 1;66(23):11314-22.
71. Nicolini FE, Basak GW, Soverini S, et al. Allogeneic stem cell transplantation for patients harboring T315I BCR-ABL mutated leukemias. *Blood.* 2011 Nov 17;118(20):5697-700.
72. Amsberg GK, Schafhausen P. Bosutinib in the management of chronic myelogenous leukemia. *Biologics.* 2013;7:115-22.
73. Cortes JE, Kantarjian H, Shah NP, et al. Ponatinib in refractory Philadelphia chromosome-positive leukemias. *N Engl J Med.* 2012 Nov 29;367(22):2075-88.
74. Cortes J, Dong-Wook K, Ibarz J, et al. A Phase 2 trial of Ponatinib in refractory Philadelphia chromosome-positive leukemias. *N Engl J Med.* 2013.
75. Breccia M, Biondo F, Latagliata R, et al. Identification of risk factors in atypical chronic myeloid leukemia. *Haematologica.* 2006 Nov;91(11):1566-8.

76. Hernandez JM, del Canizo MC, Cuneo A, et al. Clinical, hematological and cytogenetic characteristics of atypical chronic myeloid leukemia. *Ann Oncol*. 2000 Apr;11(4):441-4.
77. Jelinek J, Oki Y, Gharibyan V, et al. JAK2 mutation 1849G>T is rare in acute leukemias but can be found in CMML, Philadelphia chromosome-negative CML, and megakaryocytic leukemia. *Blood*. 2005 Nov 15;106(10):3370-3.
78. Reiter A, Invernizzi R, Cross NC, et al. Molecular basis of myelodysplastic/myeloproliferative neoplasms. *Haematologica*. 2009 Dec;94(12):1634-8.
79. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009 Aug 15;25(16):2078-9.
80. Dang L, White DW, Gross S, et al. Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature*. 2009 Dec 10;462(7274):739-44.
81. Piazza R, Valletta S, Winkelmann N, et al. Recurrent SETBP1 mutations in atypical chronic myeloid leukemia. *Nat Genet*. 2013 Jan;45(1):18-24.
82. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*. 2009 May 1;25(9):1105-11.
83. Langmead B, Trapnell C, Pop M, et al. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009;10(3):R25.
84. Robinson JT, Thorvaldsdottir H, Winckler W, et al. Integrative genomics viewer. *Nat Biotechnol*. 2011 Jan;29(1):24-6.
85. Xu G, Deng N, Zhao Z, et al. SAMMate: a GUI tool for processing short read alignments in SAM/BAM format. *Source Code Biol Med*. 2011;6(1):2.

86. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol.* 2010;11(10):R106.
87. Tarazona S, Garcia-Alcalde F, Dopazo J, et al. Differential expression in RNA-seq: a matter of depth. *Genome Res.* 2011 Dec;21(12):2213-23.
88. Piazza R, Pirola A, Spinelli R, et al. FusionAnalyser: a new graphical, event-driven tool for fusion rearrangements discovery. *Nucleic Acids Res.* 2012 Sep;40(16):e123.
89. Wang Z, Liu X, Yang BZ, et al. The Role and Challenges of Exome Sequencing in Studies of Human Diseases. *Front Genet.* 2013;4:160.
90. Davies C, Yip BH, Fernandez-Mercado M, et al. Silencing of ASXL1 impairs the granulomonocytic lineage potential of human CD34(+) progenitor cells. *Br J Haematol.* 2013 Mar;160(6):842-50.
91. Ashraf S, Noguera NI, Di Giandomenico J, et al. Rapid detection of IDH2 (R140Q and R172K) mutations in acute myeloid leukemia. *Ann Hematol.* 2013 Oct;92(10):1319-23.
92. Rakheja D, Konoplev S, Medeiros LJ, et al. IDH mutations in acute myeloid leukemia. *Hum Pathol.* Oct;43(10):1541-51.
93. Ernst T, Chase AJ, Score J, et al. Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders. *Nat Genet.* 2010 Aug;42(8):722-6.
94. Minakuchi M, Kakazu N, Gorrin-Rivas MJ, et al. Identification and characterization of SEB, a novel protein that binds to the acute undifferentiated leukemia-associated protein SET. *Eur J Biochem.* 2001 Mar;268(5):1340-51.
95. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc.* 2009;4(7):1073-81.

96. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010 Apr;7(4):248-9.
97. Papaemmanuil E, Gerstung M, Malcovati L, et al. Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood*. 2013 Sep 12.
98. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*. 2009 Jan;10(1):57-63.
99. Cristobal I, Blanco FJ, Garcia-Orti L, et al. SETBP1 overexpression is a novel leukemogenic mechanism that predicts adverse outcome in elderly patients with acute myeloid leukemia. *Blood*. 2010 Jan 21;115(3):615-25.
100. Neviani P, Santhanam R, Trotta R, et al. The tumor suppressor PP2A is functionally inactivated in blast crisis CML through the inhibitory activity of the BCR/ABL-regulated SET protein. *Cancer Cell*. 2005 Nov;8(5):355-68.
101. Puntervoll P, Linding R, Gemund C, et al. ELM server: A new resource for investigating short functional sites in modular eukaryotic proteins. *Nucleic Acids Res*. 2003 Jul 1;31(13):3625-30.
102. Druker BJ, Talpaz M, Resta DJ, et al. Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *N Engl J Med*. 2001 Apr 5;344(14):1031-7.
103. Fernandez-Mercado M, Pellagatti A, Di Genua C, et al. Mutations in SETBP1 are recurrent in myelodysplastic syndromes and often coexist with cytogenetic markers associated with disease progression. *Br J Haematol*. 2013 Oct;163(2):235-9.

104. Ott MG, Schmidt M, Schwarzwaelder K, et al. Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. *Nat Med*. 2006 Apr;12(4):401-9.
105. Panagopoulos I, Kerndrup G, Carlsen N, et al. Fusion of NUP98 and the SET binding protein 1 (SETBP1) gene in a paediatric acute T cell lymphoblastic leukaemia with t(11;18)(p15;q12). *Br J Haematol*. 2007 Jan;136(2):294-6.
106. Makishima H, Yoshida K, Nguyen N, et al. Somatic SETBP1 mutations in myeloid malignancies. *Nat Genet*. 2012 Aug;45(8):942-6.
107. Oakley K, Han Y, Vishwakarma BA, et al. Setbp1 promotes the self-renewal of murine myeloid progenitors via activation of Hoxa9 and Hoxa10. *Blood*. 2012 Jun 21;119(25):6099-108.
108. Liu X, Sun Y, Weinberg RA, et al. Ski/Sno and TGF-beta signaling. *Cytokine Growth Factor Rev*. 2001 Mar;12(1):1-8.
109. Ko JM, Lim BC, Kim KJ, et al. Distinct neurological features in a patient with Schinzel-Giedion syndrome caused by a recurrent SETBP1 mutation. *Childs Nerv Syst*. 2013 Apr;29(4):525-9.
110. Hoischen A, van Bon BW, Gilissen C, et al. De novo mutations of SETBP1 cause Schinzel-Giedion syndrome. *Nat Genet*. 2010 Jun;42(6):483-5.
111. Tartaglia M, Gelb BD. Noonan syndrome and related disorders: genetics and pathogenesis. *Annu Rev Genomics Hum Genet*. 2005;6:45-68.
112. Su AI, Wiltshire T, Batalov S, et al. A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc Natl Acad Sci U S A*. 2004 Apr 20;101(16):6062-7.

113. Mundy GR. The effects of TGF-beta on bone. *Ciba Found Symp.* 1991;157:137-43; discussion 43-51.
114. Lehman AM, McFadden D, Pugash D, et al. Schinzel-Giedion syndrome: report of splenopancreatic fusion and proposed diagnostic criteria. *Am J Med Genet A.* 2008 May 15;146A(10):1299-306.