REVIEW ARTICLE

Ferrata Storti Foundation

Integrating genetic and epigenetic factors in chronic myeloid leukemia risk assessment: toward gene expression-based biomarkers

Vaidehi Krishnan,[1,2] Dennis Dong Hwan Kim,[2,3] Timothy P. Hughes[2,4,5,6] Susan Branford[2,4,7,8] and S. Tiong Ong[1,2,9,10,11]

[1]Cancer and Stem Cell Biology Signature Research Program, Duke-NUS Medical School, Singapore, Singapore; [2]International Chronic Myeloid Leukaemia Foundation, Bexhill on Sea, UK; [3]Department of Medical Oncology and Hematology, Princess Margaret Cancer Center, University Health Network, University of Toronto, Toronto, Ontario, Canada; [4]School of Medicine, University of Adelaide, Adelaide, Australia; [5]South Australian Health & Medical Research Institute, Adelaide, Australia; [6]Department of Hematology, Royal Adelaide Hospital, Adelaide, Australia; [7]Department of Genetics and Molecular Pathology, Center for Cancer Biology, SA Pathology, Adelaide, Australia; [8]School of Pharmacy and Medical Science, University of South Australia, Adelaide, Australia; [9]Department of Haematology, Singapore General Hospital, Singapore; [10]Department of Medical Oncology, National Cancer Centre Singapore, Singapore and [11]Department of Medicine, Duke University Medical Center, Durham, NC, USA

Haematologica 2022
Volume 107(2):358-370

ABSTRACT

Cancer treatment is constantly evolving from a one-size-fits-all towards bespoke approaches for each patient. In certain solid cancers, including breast and lung, tumor genome profiling has been incorporated into therapeutic decision-making. For chronic phase chronic myeloid leukemia (CML), while tyrosine kinase inhibitor therapy is the standard treatment, current clinical scoring systems cannot accurately predict the heterogeneous treatment outcomes observed in patients. Biomarkers capable of segregating patients according to outcome at diagnosis are needed to improve management, and facilitate enrollment in clinical trials seeking to prevent blast crisis transformation and improve the depth of molecular responses. To this end, gene expression (GE) profiling studies have evaluated whether GE signatures at diagnosis are clinically informative. Patient material from a variety of sources has been profiled using microarrays, RNA sequencing and, more recently, single-cell RNA sequencing. However, differences in the cell types profiled, the technologies used, and the inherent complexities associated with the interpretation of genomic data pose challenges in distilling GE datasets into biomarkers with clinical utility. The goal of this paper is to review previous studies evaluating GE profiling in CML, and explore their potential as risk assessment tools for individualized CML treatment. We also review the contribution that acquired mutations, including those seen in clonal hematopoiesis, make to GE profiles, and how a model integrating contributions of genetic and epigenetic factors in resistance to tyrosine kinase inhibitors and blast crisis transformation can define a route to GE-based biomarkers. Finally, we outline a four-stage approach for the development of GE-based biomarkers in CML.

Correspondence:

S. TIONG ONG
sintiong.ong@duke-nus.edu.sg

Received: July 16, 2021.

Accepted: September 28, 2021.

Pre-published: October 7 2021.

https://doi.org/10.3324/haematol.2021.279317

Introduction

Chronic myeloid leukemia (CML) is a clonal disorder of the hematopoietic stem cell compartment defined and driven by the *BCR-ABL1* gene rearrangement and the tyrosine kinase it encodes.[1] Clinically, it is accompanied by an expansion of mostly myelo-erythroid progenitors that maintain the ability to differentiate terminally into neutrophils. Prior to the introduction of ABL1 tyrosine kinase inhibitors (TKI), most patients would progress to a terminal blast crisis (BC) stage marked by the acquisition of additional genetic abnormalities within an average of 5-7 years.[2] In this stage, the clinico-pathological features were the inexorable accumulation of either myeloid or lymphoid progenitors that had acquired aber-

rant self-renewal properties, broad resistance to cytotoxic therapies, and eventual patient demise from bone marrow failure.[2] The arrival of TKI at the turn of the century resulted in remarkable responses, such that most individuals treated in chronic phase (CP) CML can expect to achieve near-normal life expectancies.[3] Nevertheless, CML-related deaths are still reported, mainly due to resistance and progression to BC, especially in the first few years of treatment.[4]

Current treatment aims and features of an ideal biomarker

Current therapeutic aims are directed at achieving sufficiently deep molecular responses that the risks of BC transformation are effectively negligible and, in the longer-term, increasing the rates of treatment-free remission.[5-7] Clinical guidelines toward achieving deep molecular responses have been reviewed elsewhere,[8] and at their core, prescribe the measurement of *BCR-ABL1* transcript levels using the International Scale (IS) every 3 months as a readout of the depth of the response to TKI. In turn, the depth of TKI response serves as a critical biomarker guiding patient management and prognostication (Figure 1).

Given current treatment goals, an ideal biomarker would accurately predict patients who will achieve a deep molecular response with first-line TKI, or require a switch to alternative therapy, and, among those who achieve a deep molecular response, those who will be able to stop TKI successfully (Figure 1). The biomarker would be informative from the time of diagnosis and prior to TKI initiation, since this would enable early stratification of patients for therapy with a first-generation *versus* a second/third-generation TKI, allosteric BCR-ABL1 inhibitor, a clinical trial, or preparation for allogeneic transplantation. Additionally, among patients who meet the criteria for stopping TKI therapy, the ideal biomarker would identify additional therapies that would enhance treatment-free remissions. Finally, gene expression (GE)-based biomarkers should be clinically robust, and widely available among centers and regions in both low and high Human Development Index countries.[9]

Why gene expression-based biomarkers?

Contributions from genetic and epigenetic mediators to TKI resistance and BC transformation are well documented,[10-14] and it is axiomatic that genetic or epigenetic factors mediating these outcomes will contribute to a cell's GE signature. Accordingly, GE signatures offer a molecular profile that integrates risk factors encoded by both mutations and epigenetic states. However, faithfully extracting and interpreting GE-based information in clinical settings is challenging. Barriers to adoption include technical limitations, logistical factors, as well as differences in study design and data analysis, and are described below.

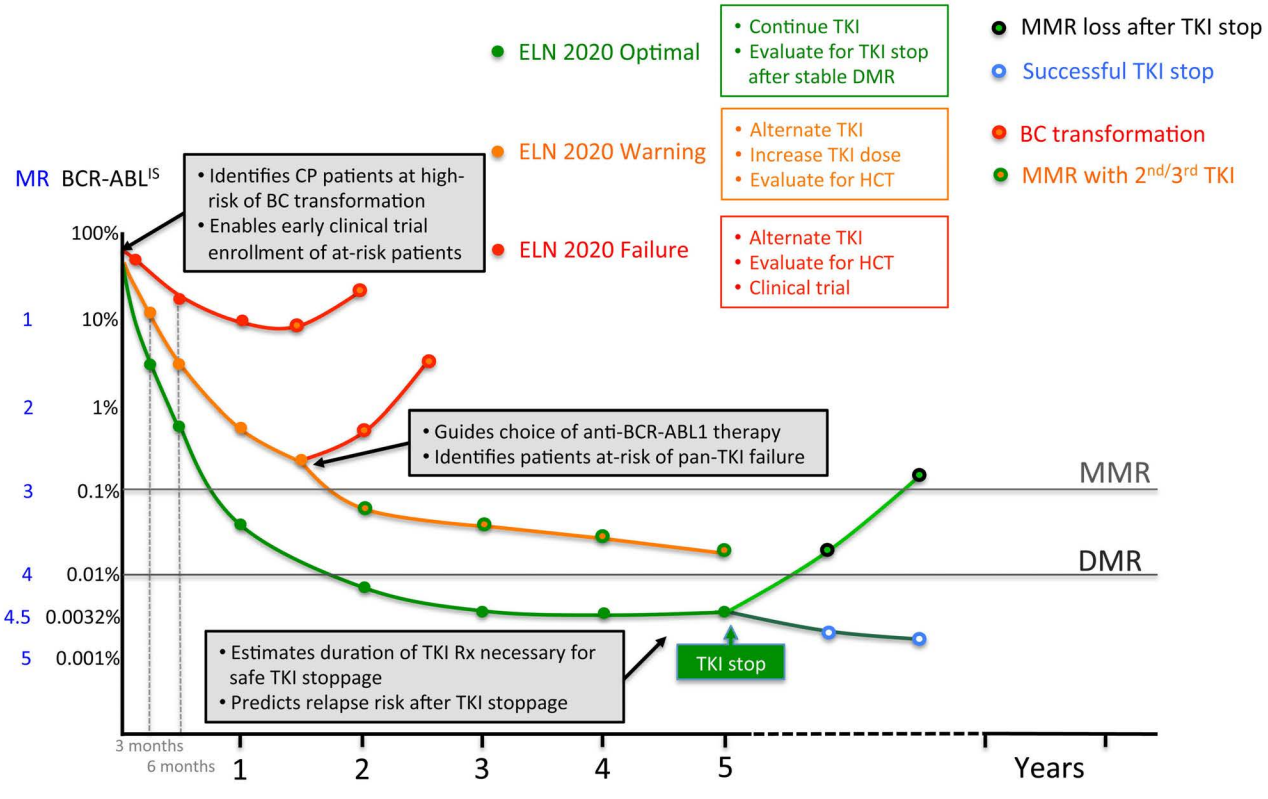

Figure 1. Features of an ideal chronic myeloid leukemia biomarker. Curves indicate changes in *BCR-ABL1* transcript levels, meausred using the International Scale (BCR-ABL1^{IS}), following initation of tyrosine kinase inhibitor (TKI) therapy in patients with chronic phase (CP) chronic myeloid leukemia (CML). The corresponding molecular response (MR) value is provided next to the BCR-ABL^{IS} value. Green, orange, and red curves are representative of patients in European LeukemiaNet 2020 'optimal', 'warning', and 'failure' cateogories respectively. Major (MMR) and deep (DMR) molecular remissions are defined as 0.1% (MR 3) and 0.01% (MR4) BCR-ABL^{IS}, respectively. Green, orange, and yellow bullet points indicate guidelines for each category of response. Gray boxes describe predictive capabilities of an ideal biomarker. HCT: hematopoietic stem cell transplant.

Despite the barriers, recent advances in technological and computational platforms are enabling the interrogation of patient samples on an unprecedented scale, and are being translated into robust technical assays on patient material that are reproducible in clinical laboratories.[15] Such advances may eventually result in the identification of pretreatment biomarkers that not only predict TKI resistance but suggest alternative non-BCR-ABL1-targeting therapies to pre-empt the emergence of clinical resistance. Accordingly, it is timely to review the results of GE studies using primary patient material annotated for clinical outcomes, and assess how genetic and epigenetic factors associated with treatment outcome contribute to GE signatures. In doing so, it is also important to develop models incorporating the interplay between genetic and epigenetic factors, and determine how best to use the resulting GE outputs to understand and predict CML drug resistance and transformation. Finally, it is incumbent on the CML community to outline the practical steps needed for the clinical development of GE-based biomarkers in CML.

Gene expression signatures associated with resistance to tyrosine kinase inhibitors

Since the beginning of the TKI era, a variety of diagnostic material from CP patients has been used to discover TKI-resistance GE signatures (Table 1). Here, we review the key conclusions from these studies.

Gene expression using peripheral blood

In the earliest research by Kaneta *et al.*[16] and McLean *et al.*[17] microarray studies were conducted on blood from imatinib responders and non-responders. Apart from *CBLB*, which was downregulated in responders, there was no overlap between the two datasets. De Lavallade *et al.* conducted microarray studies on peripheral blood mononuclear cells to identify a 105-gene set that was enriched in imatinib non-responders, comprising mainly genes in cell cycle and DNA repair pathways.[18] However, the GE signature could be validated only in an imatinib-treated cohort but not in a cohort treated with interferon-α. As a targeted approach, the expression of 21 genes associated with TKI responses and disease progression was studied by Zhang *et al.*[19] Increased *PTGS1* expression was the only gene that differentiated primary imatinib-resistant patients from responders, while 15 genes distinguished CP from BC. Twelve genes distinguished imatinib-responsive from secondary imatinib-resistant CML without *BCR-ABL1* mutations, of which *LYN*, *JAK2*, *PTPN22* and *CEBPA* downregulation was shared with BC samples. The study concluded that at least some features of secondary imatinib resistance overlap with BC transformation.

More recently, Kok *et al.* conducted microarray-based analysis on diagnostic blood from 96 CP patients from the TIDEL-II trial to predict failure of early molecular response,[20] which correlates with inferior long-term outcomes.[21,22] Three hundred sixty-five differentially expressed genes were identified which were enriched for 'cell cycle' and 'stemness' (*MYC*, *HOXA9*, β-catenin) but depleted for 'immune-response' categories in the group with early molecular response failure. A binary classification model was built to predict early molecular response failure based on 17 genes and the signature was validated in an independent cohort. Of these, eight genes *IGFBP2*, *SRSF11*, *BAX*, *CDKN1B*, *BNIP3L*, *FZD7*, *PRSS57*, and *RPS28* intersected with findings of previous CML TKI-resistance and progression studies. This study demonstrated that GE information from diagnostic samples could predict events long in the future, including major molecular response (MMR) at 24 months, MR4.5 at 5 years, and BC transformation.

Gene expression using bone marrow

Independently, a series of studies used unselected bone marrow for comparisons of GE between groups of patients with different treatment responses. Frank *et al.* identified a 128 GE signature associated with imatinib resistance, specifically in an interferon-α pre-treated cohort. Differentially expressed genes were involved in apoptosis (*CASP9*, *TRAP1*), DNA repair (*MSH3*, *DDB2*), oxidative stress protection (*GSS*, *PON2*, *VNN1*) and centrosomes (*ID1*).[23] Villuendas *et al.*[24] identified 46 differentially expressed genes of which a six-gene prediction score (*BIRC4*, *FZD7*, *IKBKB*, *IL-7R*, *TNC*, *VWF*) that correlated with imatinib resistance after interferon-α failure developed. Differentially expressed genes were involved in cell adhesion (*TNC* and *SCAM-1*), drug metabolism (*COX1* or *PTGS1*), protein tyrosine kinases (*MKNK1*), and phosphatases (*BTK* and *PTPN22*). Notably, the MKNK1/2 kinases have been shown by two independent groups to be involved in BC transformation.[25,26] In contrast to the prior studies, Crossman *et al.* found no differentially expressed genes between the imatinib responder categories. The use of mixed peripheral blood and bone marrow samples, unselected white blood cells and a heterogeneous cohort of patients in late CP and heavily pre-treated, were suggested as potential reasons for the negative results.[27] The important conclusion was that GE comparisons should be made on purified CD34$^+$ cells. Indeed, in a meta-analysis comparing six published GE studies in CML, *DDX11*, *MSH5*, and *RAB11FIP3* were the only genes coincident between any two of the studies.[28] The small differences in differential GE between responder groups, different GE platforms, different statistical methods and different sources of cells profiled were suggested reasons for the poor intersection. The disappointing results from unselected peripheral blood and bone marrow provided the impetus to isolate and study CD34$^+$ fractions.

Gene expression using CD34$^+$ cells

McWeeney *et al.* were the first group to use CD34$^+$ cells from diagnostic bone marrow.[11] Cell adhesion genes were upregulated in imatinib-resistant patients suggesting that CD34$^+$ cells may establish more adhesive interactions with the bone marrow milieu. The enrichment for β-catenin binding targets suggested activated Wnt/β-catenin signaling in imatinib-resistant patients, a feature shared with CD34$^+$ progenitors from BC.[26,29] The authors concluded that primary resistance to imatinib might reflect more advanced disease progression. A 75-probe minimal gene classifier predicted 88% of responders and 83% of non-responders in a validation cohort. Importantly, the authors of this paper compared their GE signatures to those predicting early BC transformation, as discussed below, and provided an important resource for validation and comparison of other CD34$^+$-based GE datasets.

Table 1. Gene expression profiling studies comparing responders and non-responders to tyrosine kinase inhibitors.

	Kaneta *et al.*, 2002	McLean *et al.*, 2004	Crossman *et al.*, 2004	Villuendas *et al.*, 2006	Frank *et al.*, 2006	McWeeney *et al.*, 2009	Zhang *et al.*, 2009	de Lavallade *et al.*, 2010	Kok *et al.*, 2019
Stage & numbers	CP 18; AP 2; BC 2	CP 66	CP 29, included patients previously on IFN-α	CP 32, 12 validation	CP 23 R; 11 NR	CP 12 R; 24 NR (discovery); CP17 R 6 NR (validation)	CP 63; AP 5; secondary TKI-R 29; BC 27	CP 15	CP 96 (discovery); CP 88 (validation); CP 132 (nilotinib Rx).
Time sample taken	Diagnostic	Diagnostic	Prior to TKI, but could have been on IFN-α	Diagnostic BM	Diagnostic PB & BM	Diagnostic PB & BM	Diagnostic blood	Diagnostic PB	Diagnostic PB
Unselected/ $CD34^+$ cells; PB/BM	Unselected; PB & MNC	Unselected; whole blood	Unselected; total WBC from PB & BM	Unselected; BM	Unselected; total WBC from BM & PB	$CD34^+$ selected BM MNC & $CD34^+$ PB MNC in validation group	Unselected; total WBC	Unselected; MNC	Unselected; MNC
Platform	cDNA Microarray	Microarray (Affymetrix HG_U95Av2)	Microarray (Affymetrix HG_U95Av2)	Microarray (CNIO OncoChip)	Microarray (Affymetrix HG-U133A)	Microarray (Affymetrix HG-U133 Plus 2.0)	TaqMan LDA	Microarray (Affymetrix HG-U133 Plus 2.0)	Microarray (Illumina HT-12v4) TaqMan LDA
N. of DEG	79	55	-	46	128	885	21	105	365
Time of predicted event	R=MCyR (<35% Ph^+); NR= >65% Ph^+ at 5 months	R=CCyR (0% Ph^+); NR= >65% Ph^+ at 12 months	R=CCyR within 9 months; NR= >35% Ph^+ after 1 year	MCyR at 12 months	R=MCyR (≤35% Ph^+); NR= ≥35% Ph^+ at 12 months	R=CCyR at 12 months; NR= >66% Ph^+ at 12 months	CCyR at 12 months	CCyR at 12 months; NR (failed to achieve any cytogenetic response)	EMR at 3 months
Biological insights	First evidence that GE profiles can predict sensitivity to imatinib	Predictive genes enriched for cell adhesion, mitogenic signaling, apoptosis	GE comparisons should be made on purified $CD34^+$ cells	Predictive genes associated with Wnt signaling, cell adhesion, NK-κB, apoptosis, DNA repair	Predictive genes enriched for transcriptional regulation of apoptosis, oxidative stress, DNA repair, centrosomal genes	Predictive genes enriched for cell adhesion and targets of the Wnt/β-catenin pathway	Predictive genes involved in TKI influx/efflux, BC progression, BCR-ABL1 signaling; Secondary TKI-R genes similar to BC genes but not primary TKI-R	Predictive genes enriched for DNA repair by recombination	GSEA indicated genes associated with poorer outcome enriched for cell cycle, stem cell function, & depleted for immune function.
Comments	79 DEG were identified. 15 or 30 genes were used to develop a prediction score to separate TKI-responders from non-responders	31 genes were used to develop a classifier to separate TKI-responders from non-responders.	No DEG were identified between TKI responders and non-responders	A 6-gene prediction model was constructed which could predict MCyR at 12 months	A 128-gene predictor of primary cytogenetic resistance to imatinib was identified	A 75-probe set classifier that separated the responder groups. PPV 87.7%; NPV 73.7% PPV 94.4%; NPV 75%. $CD34^+$ cell selection & microarray analysis possible, successful in 71% of patients. Predictive genes overlapped with three independent datasets for BC genes (Zheng *et al.*, 2006), genes prediciting early BC transformation (Yong *et al.*, 2006), PRC target genes in BC (Ko *et al.*, 2020).	15 genes distinguished CP from BC. 12 genes distinguished between secondary TKI-R *vs.* optimal responders. *PTGS1* predicted primary TKI-R	Identified a set of genes whose expression was differentially regulated in patients resistant to imatinib	A binary classification model based on 17 genes. HR-GES: 77% failure, but missed 2/9. LR-GES: 95% did well, but missed 4/79. 64% sensitivity; 97% specificity. HR-GES had lower rate of EMR failure with nilotinib

CP: chronic phase; AP: accelerated phase; BC: blast crisis; R: responder; NR: non-responder; TKI-R: resistance to tyrosine kinase inhibitors; Rx: treatment; TKI: tyrosine kinase inhibitor; IFN-α: interferon-alpha; BM: bone marrow; PB: peripheral blood; PBMC: peripheral blood mononuclear cells; MNC: mononuclear cells; WBC: white blood cells; N.; number; DEG: differentially expressed genes; MCyR: major cytogenetic response; CCyR: complete cytogenetic response; Ph^+: Philadelphia chromosome-positive; EMR: early molecular response; GE: gene expression; GSEA: gene set enrichment analysis; DEG: differentially expressed genes; PPV: positive predictive value; NPV: negative predictive value; HR-GES: high-risk gene expression signature; LR-GES: low-risk gene expression signature.

Single-cell-based gene expression analysis

Recent advances in single-cell analysis have enabled novel GE-based insights on the roles of tumor cell heterogeneity and clonal evolution under the selective pressure of therapeutics, with obvious applications in biomarker development.[30] Leukemia stem cell (LSC) heterogeneity was characterized by Warfvinge *et al.* by combining high-throughput immunophenotyping with single-cell GE profiling with a defined panel of genes.[31] LSC sub-fractions with more primitive and quiescent signatures had a higher persistence after TKI therapy with the most TKI-insensitive population identified as $Lin^{-}CD34^{+}CD38^{low/-}CD45RA^{-}cKIT^{-}CD26^{+}$ stem cells. Giustacchini *et al.* used the Smart-seq2 platform to combine single-cell RNA-sequencing analysis with *BCR-ABL1* transcript detection using purified stem cells. A sub-population of $BCR\text{-}ABL1^{+}$ quiescent stem cells enriched for hematopoietic stem cell (HSC) signatures was found to persist during TKI therapy.[32] Intriguingly, the $BCR\text{-}ABL1^{-}$ cells in CML patients were enriched for inflammatory, tumor growth factor-β and tumor necrosis factor-α hallmarks and discriminated between the TKI-responder groups.

In addition to LSC-derived signatures, the GE signature of immune cells can be equally instructive. For example, plasmacytoid dendritic cells, the major producers of interferon-α *in vivo*, promoted resistance to nilotinib in CML patients.[33] These studies imply that the cytokines released by immune cells in the bone marrow microenvironment, and the transcriptomic changes that they bring about on the LSC, may activate cytokine-dependent TKI resistance programs.[34] Together, these single-cell studies demonstrate that GE signatures within malignant and non-malignant compartments in CML are prognostically informative. We anticipate the discovery of additional biomarkers among discrete cell types which have the potential to be assayed by platforms available in standard pathology laboratories, e.g., by flow cytometry or immunohistochemistry.

Gene expression signatures associated with blast crisis progression

Transcriptomic comparisons between the CP and BC stages have uncovered progression-related signatures that can herald BC transformation (Table 2). In the pre-TKI era, the time to BC transformation from CP varied between patients, and to understand this difference, Yong *et al.* compared $CD34^{+}$ cells from leukapheresis samples provided by patients who progressed to BC within 3 years (aggressive leukemia) or after more than 7 years (indolent leukemia) following diagnosis.[13] The study identified that lower *CD7* with higher *PR3* and *ELA2* expression at diagnosis was associated with longer survival. Intriguingly, when the GE signatures identified by Yong *et al.* and McWeeney *et al.* were compared, a significant overlap was found.[11] This important study demonstrated that biological processes associated with TKI resistance and early BC transformation overlapped, and that $CD34^{+}$ cells from different sources (bone marrow *vs*. peripheral blood) contained this information.

In the post-TKI era, a landmark study by Radich *et al.* identified distinct transcriptional programs during BC progression.[25] About 3,000 genes were associated with the BC stage with a dysregulated WNT/β-catenin pathway, decreased *Jun B* and *FOS*, and higher *PRAME* expression. The Radich dataset was subsequently used to compute a six-gene signature comprising *NOB1*, *DDX47*, *IGSF2*, *LTB4R*, *SCARB1*, and *SLC25A3* to predict progression.[35] Independently, Zheng *et al.* isolated $CD34^{+}$ cells and identified 34 differentially expressed genes as cells transited from CP to BC. Among the misregulated genes, *SOCS2* and *CD52* were downregulated while HLA-related genes were overexpressed in BC.[36]

To understand the biological mechanisms underlying GE changes in TKI resistance and BC, a recent study tested the hypothesis that prognostically important genes were enriched for targets of the polycomb repressive complex (PRC; see below).[10] Importantly, target genes of PRC-asso-

Table 2. Gene expression profiling studies comparing chronic phase and acute phase.

	Yong *et al.*, 2006	Radich *et al.*, 2006	Zheng *et al.*, 2006	Oehler *et al.*, 2009	Ko *et al.*, 2020
Stage & numbers	CP 68	CP 42; AP 17; BC 32	CP 11; BC 9	CP 42; AP 17; BC 34	CP 16; MBC 13; LBC 5
Unselected/$CD34^{+}$; PBMC/BM	$CD34^{+}$; PBMC	Unselected; BM	$CD34^{+}$; PBMC	Unselected; BM	$CD34^{+}$; PBMC
Platform	Microarray	Microarray	Microarray	Microarray	Microarray
N. of genes	20	3000+	114	6	431 Upregulated LBC 522 downregulated
Comments	Identifies early (≤3 years) *vs.* late (≥7 years) BC transformation. Low CD7 & high PR-3 predicts higher OS.	Identifies TKI-R in CP (had BC-like signature)	Genes that distinguish CP and BC	Discriminates between early & late CP	Identifies a core BC gene expression signature common to MBC and LBC. PRC-driven transcriptional reprogramming is enriched for poor prognostic genes in CP in the $CD34^{+}$ datasets of Yong *et al.* (2006) and McWeeney *et al.* (2009).

CP: chronic phase; AP: accelerated phase; BC: blast crisis; MBC: myeloid blast crisis; LBC: lymphoid blast crisis; PBMC: peripheral blood mononuclear cells; BM: bone marrow; PR-3: proteinase-3: OS: overall survival; TKI-R: resistance to tyrosine kinase inhibitors; PRC: polycomb repressive complex.

ciated silencing in BC progression were enriched for downregulated genes identified in the datasets of both McWeeney *et al.* and Yong *et al.*[10] The cross-validation of these three independent datasets suggests important lessons for the development of GE-based risk assessment: (i) the discovery of reproducible GE-based biomarkers is possible when homogeneous CD34$^+$ populations are used; (ii) the processes of TKI resistance and BC transformation are biologically convergent despite genetic heterogeneity:[10] and (iii). PRC-regulated processes contribute to silencing of prognostically informative genes.

Contribution of somatic mutations to gene expression signatures

Recent reviews have described the range and frequency of specific genetic mutations in patients who developed TKI resistance and/or BC.[37] For many of these genes, there is strong preclinical information indicating that their associated mutations contribute to or are even sufficient to produce TKI resistance or transformation phenotypes (summarized in Table 3).[38-48] These studies imply that genetic mutations alter GE profiles, and here we review their contributions to GE changes in CML since these changes may represent useful GE-based biomarkers.

For *RUNX1* mutations, the Mustjoki group identified an accompanying GE signature in BC samples.[49] They found that *RUNX1* mutations were associated with the upregulation of stemness, B-cell markers, interferon and immune signaling and transcription factors regulating plasmacytoid dendritic cell development.

In analogous work, the overexpression of an *IKZF1* dominant-negative mutant in CD34$^+$ cells from CP patients increased STAT5 expression, a pathway associated with imatinib resistance,[50] and enhanced transformation.[51] *RAG* expression status was recently assessed in diagnostic samples, given the role of RAG recombination as a mediator of *IKZF1* deletions.[52] Notably, *RAG1/2* and *DNTT* upregulation at diagnosis suggested imminent lymphoid BC transformation within 12 months (8/8 patients), demonstrating that GE signatures can reliably predict transformation.

Despite limited functional interrogation of *ASXL1* using CML patient material, insertion sites within *ASXL1* promoted BC progression in a CP mouse model subjected to transposition-based mutagenesis.[53] Transgenic expression of truncated protein ASXL1$^{aa1\text{-}587}$ in mice increased HSC self-renewal, and Brd4 occupancy and chromatin accessibility around genes required for stemness, and predisposed mice to myeloid malignancies.[54] However, the clinical relevance of diagnostic *ASXL1* mutations is still unclear because some patients with *ASXL1* variants at diagnosis can achieve a MMR after TKI therapy.[14] Furthermore, *ASXL1* mutations frequently disappeared when monitored in the long-term during TKI therapy (*personal observation by Dr. Dennis Kim*). Meanwhile, direct evidence for contributions of other mutations to CML GE signatures is currently lacking, and we have to infer them from studies in other malignancies (Table 3).

Lessons from clonal hematopoiesis

Clonal hematopoiesis is the clinical phenomenon by which populations of hematopoietic cells expand and carry a somatic mutation that is at least 2% of the variant allele fraction.[55] The common genes comprise *DNMT3A*, *TET2*, and *ASXL1*, and others also found in CML individuals, including *RUNX1*, *BCORL1*, and *TP53*.[56] Individuals with clonal hematopoiesis are at increased risk of developing hematologic malignancies, and it is therefore pertinent to ask whether clonal hematopoiesis-related mutations also confer increased risk of TKI resistance or progression. A study by Kim *et al.* has highlighted important features of clonal hematopoiesis-related mutations in CML.[57] Firstly, they may occur in a non-Philadelphia chromosome-positive clone and predate the development of CML, and are unrelated to the CML clone. Secondly, even when a specific mutation occurs in the Philadelphia chromosome-positive clone, it only confers a relative risk of TKI resistance or progression. Indeed, patients with *RUNX1* mutations have been documented to achieve MMR (*personal observation, Dr. Dennis Kim*). Nevertheless, Kim *et al.* concluded that mutations in genes regulating epigenetic function (*TET2*, *ASXL1* among them) were associated with a higher risk of inferior TKI responses.

There are also strong preclinical data indicating that clonal hematopoiesis-related mutations result in subtle but important changes in GE in HSC. For example, *Dnmt3a*-deficient HSC show a loss of DNA methylation in regions enriched for self-renewal genes such as *Meis1*, *Evi1* and *HoxA9*.[58] In Tet2-deficient mice, the loss of DNA demethylation is accompanied by an expansion of the stem and progenitor cell compartments, and eventual myeloproliferation.[45] In *ASXL1*-deficient mice, an increase in self-renewal capacity of stem cells is observed, through the loss of PRC1-mediated gene repression.[59] Another interesting aspect of hematopoietic stem and progenitor cells harboring inactivating mutations of *DNMT3A* and *TET2* is that they both led to increased cytokine production in peripheral myeloid cells, including interleukin-6 and interleukin-1β.[60,61] Furthermore, mutations associated with clonal hematopoiesis are frequently found in monocytes, granulocytes, and natural killer cells compared to B or T cells, suggesting that their effects may also be manifest in multiple differentiated cell types within the hematopoietic compartment.[62] Together, these observations are relevant to the search for prognostic GE signatures in CML for the following reasons: (i) increased inflammation and cytokine production is associated with LSC persistence,[32] and disease progression;[10,63] (ii) prognostic GE changes may be found in both CD34$^+$ and CD34$^-$ fractions of peripheral blood or bone marrow mononuclear cells; and (iii) changes in natural killer cell function and number may predict treatment-free remissions, and presumably contain informative natural killer cell GE signatures.[64-66]

Epigenetic contributions to gene expression signatures

Polycomb repressive complex-associated gene expression changes

Among the most well studied epigenetic complexes in CML are the polycomb group (PcG) proteins.[67] The polycomb group proteins assemble into two complexes, PRC2 and PRC1, which modify histones through repressive H3K27 trimethylation (H3K27me3) and H2AK119

monoubiquitination (H2AK119ub1),respectively, and in general repress gene expression.[68]

From a mechanistic standpoint, the most commonly occurring mutations in CML appear to converge in their ability to interact with and function in conjunction with the polycomb group proteins. ASXL1 functions in transcriptional repression through its interaction with PRC2 and BAP1.[69] BCORL1 is a transcriptional co-repressor that interacts with PCGF1, a core complex of the PRC1.1 complex.[70] The RUNX1-CBF-β heterodimer mediates transcription by binding to *RUNX* sites, but also represses transcription by interacting and recruiting BMI1 of the PRC1 complex to target sites.[71] IKZF1 regulates transcription by interacting with repressive epigenetic complexes such as HDAC1, HDAC2, CHD3, CHD4, and SWI/SNF complex, and also recruits PRC2 to target gene loci in T cells.[72] Thus, while the commonly mutated genes in CML have their own exclusive roles in transcriptional regulation, they also share a striking commonality as modulators of the PRC. Whether mutated variants of *RUNX1*, *ASXL1*, *IKZF1*, and *BCORL1* drive aberrant PRC recruitment and GE in CML remains to be determined.

In this respect, a recent study determined that lymphoid and myeloid BC transcriptomes are highly congruent, and that both undergo PRC-driven epigenetic reprogramming towards a convergent transcriptomic state.[10] PRC-dependent epigenetic reprogramming was attributed to gain- and loss-of-function mutations in members of the PRC1 and PRC2 complexes, respectively. Of these, ongoing BMI1/PRC1 activity contributes to maintaining the BC transcriptome, while EZH2/PRC2-binding was instructional for DNA hypermethylation-dependent gene repression. Importantly, the integrative model proposed by Ko *et al.* suggests that enrichment for PRC-dependent GE signatures at diagnosis can predict disease transformation and TKI resistance, as highlighted above.[10]

We also note that dysregulated regulation of PRC has been identified as a key feature of TKI-resistant LSC in CP. *EZH2* expression was higher in CML LSC than in normal HSC, and CML LSC have a stronger dependence on

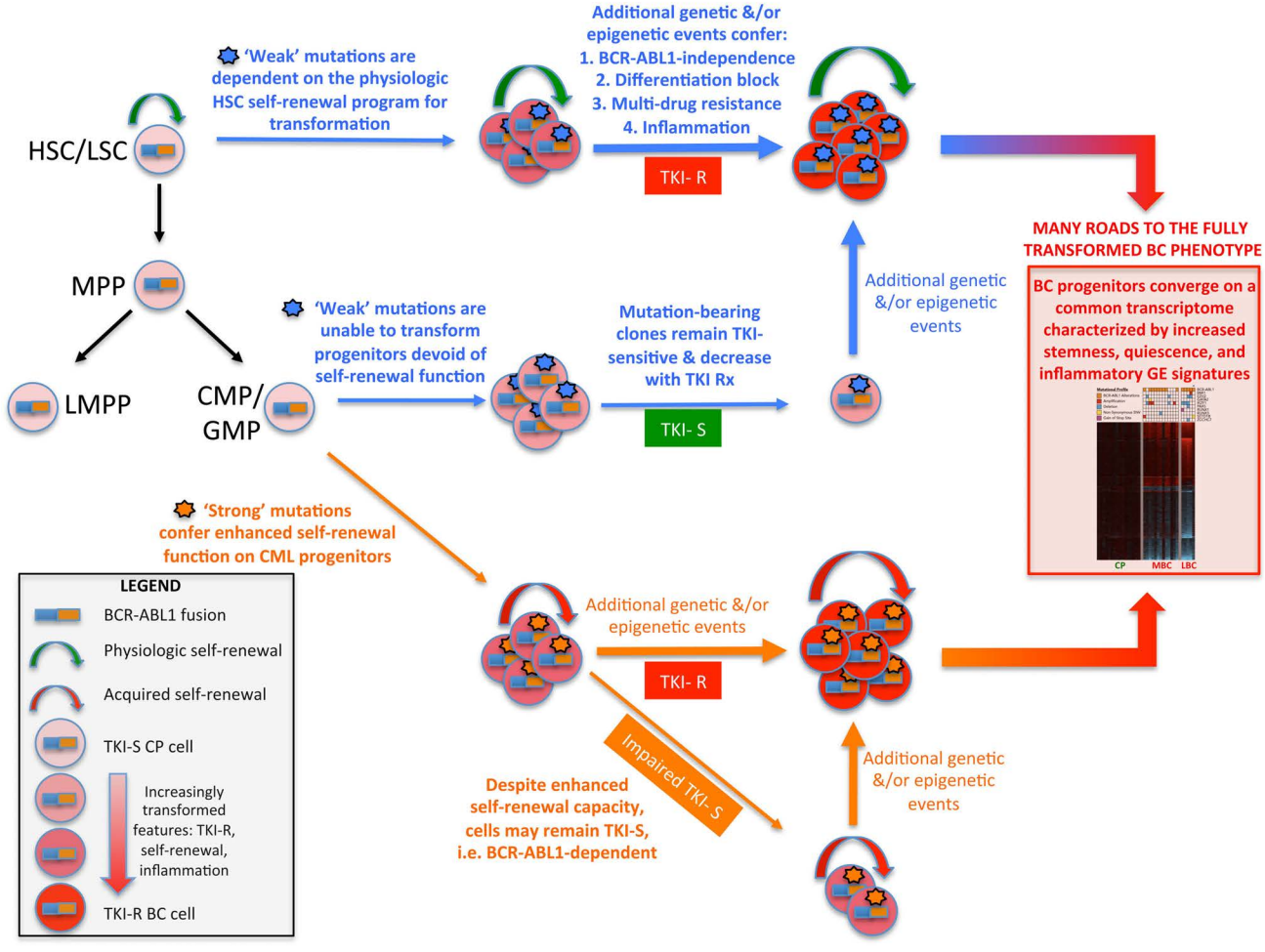

Figure 2. Diagrammatic representation of the 'seed and soil' model of chronic myeloid leukemia. The model proposes that both acquired mutations and the cell state of the mutation-acquiring cell contribute to the process of full transformation to blast crisis (BC). A 'strong' mutation is defined as being able to confer self-renewal function on a progenitor cell that does not possess inherent self-renewal capacity. A 'weak' mutation is unable to confer self-renewal function and can only transform a cell with native self-renewal ability, i.e., a stem cell. For both strong and weak mutations, it is likely that additional genetic and epigenetic events are necessary to confer the full suite of features required for BC transformation. The model is also based on the recent finding that BC progenitors which harbor different somatic mutations share a common or core transcriptome enriched for stemness, quiescence, and inflammatory gene expression signatures.[20] HSC: hematopoietic stem cell; LSC: leukemia stem cell; MPP: multipotential progenitor; LMPP: lymphoid-primed multipotent progenitor; CMP: common myeloid progenitor; GMP: granulocyte-macrophage progenitor; TKI-S/R: tyrosine kinase inhibitor-sensitive/resistant cells. Rx: treatment; CP: chronic phase; BC: blast crisis; CML: chronic myeloid leukemia; GE: gene expression.

Table 3. **Functional effects of frequently mutated genes in blast crisis.**

	RUNX1	IKZF1	ASXL1	BCORL1	GATA2	TET2	DNMT3A
Frequency: diagnosis/ progression (%)	2.6/18.3	6.1/16.0	9.7/15.1	0.9/8.6	0/8.4	0.9/6.7	2.3/4.5
Function	Transcription factor	Transcription factor	Transcriptional repression	Transcriptional co-repressor	Transcription factor	Methylcytosine dioxygenase	DNA methylation
Mode of action and interactions	DNA binding via RUNT domain.	DNA binding via zinc finger domain.	Regulator of H3K27me3 & H2AK119ub1 marks	Represses transcription by binding to class II HDAC & CTBP1	DNA binding via zinc finger domain	Conversion of 5-methylcytosine to 5-hydroxymethylcytosine	Transfer of methyl group to cytosines on DNA
Interactions with other complexes	P300, CBP, PRC1, NuRD, SWI/SNF, MLL/TrxG	HDAC1, HDAC2, CHD3, CHD4, PRC2, CtBP1, SWI/SNF	PRC2 , BAP1 complexes	PCGF1, the core PRC1.1 component	FOG1 through N-terminal Zinc finger domain	2-HG, vitamin C, OGT, WT1, VPRBP, IDAX	PRC2, EVI1, ISGF3, AP2a, ZEB1, HDAC1
Aberration in CML	RUNT domain mutations deletions and fusions	Deletions- exon Δ4–7 (IK6), exon Δ2–7	Majority are frameshift and nonsense mutations in exon 12	Frameshift, nonsense mutations	Zinc finger domain variants	Missense, nonsense and frameshift mutations TET2 mutations may be CHIP-related or a part of the Ph^+ clone	DNMT3A mutations are mostly CHIP mutations since they are also present in the Ph^- clone
Effects on gene expression	Up: Interferon signaling, immune molecules, pDC- TF. Down: DNA repair	Up: JAK-STAT signaling, self-renewal genes. Down: B-cell lineage and DNA repair genes	Increased Brd4 occupancy and chromatin accessibility around genes	Transcriptional repressor of E-cadherin. Other targets unknown	CML L359V mutant inhibits transactivation by PU.1. GATA2 MDS and AML mutants have altered transactivation activity	Impaired 5-methylcytosine hydroxylation and decreased methylation at CpG sites in myeloid cancers with mutant *TET2*	DNMT3A-deficient HSC show loss of DNA methylation at the edge of hypomethylated canyon regions enriched for self-renewal genes such as *MEIS1*, *EVI1*, *HOXA9*
Effect of mutant protein/ gene knock out in vitro	RUNX1 H78Q orV91 fs-ter94 in 32D-BCR-ABL1 model blocked differentiation	*IK6* expression in $CD34^+$ cells isolated from CP-CML patients enhances their *in vitro* expansion	Truncated *ASXL1* increased proliferation, and decreased differentiation along megakaryocyte and erythroid lineages	*BCORL1* depletion increased the re-plating capacity of Runx1-depleted Lin^- cells	MDS and AML GATA2 mutants inhibit differentiation and apoptosis	*TET2* silencing in human $CD34^+$ cells increased the monocytic lineage at the expense of erythroid and lymphoid lineages	Nearly a third of CHIP-related DNMT3A mutations reduce protein stability
Effect of mutant protein/ gene knock out *in vivo*	RUNX1 H78Q or V91fs-ter94 mutants induced a BC or accelerated phase-like phenotype in mice	Ikaros DNA binding domain inactivation in early pre-B cells leads to ALL	AML, MPN, MDS-like diseases	Effect of CML *BCORL1* variants unknown	*GATA2* deficiency has been recognized as a major MDS predisposition syndrome in humans	Conditional *TET2* loss in the hematopoietic compartment leads to increased stem cell self-renewal	*Dnmt3a* ablation in HSC predisposes mice to develop a spectrum of myeloid and lymphoid malignancies
Effect on CML variant studied	Yes	Yes	Yes	No	No	No	No
References	Zhao *et al.*, 2012 Awad *et al.*, 2020 Branford *et al.*, 2018	Joshi *et al.*, 2014 Beer *et al.*, 2015	Yang *et al.*, 2018 Katoh, 2013 Balasubramani *et al.*, 2015	Pagan *et al.*, 2007 Wong *et al.*, 2016	Zhang *et al.*, 2009 Branford *et al.*, 2018	Kim *et al.*, 2017 Pronier *et al.*, 2011 Crusio *et al.*, 2011	Mayle *et al.*, 2015 Hervouet *et al.*, 2018 Huang *et al.*, 2018 Kim *et al.*, 2017 Branford *et al.*, 2018

HDAC: histone deacetylases; CML: chronic myeloid leukemia; CHIP: clonal hematopoiesis of indeterminate potential; MDS: myelodysplastic syndrome; AML: acute myeloid leukemia; HSC: hematopoietic stem cell; CP: chronic phase; ALL: acute lymphoblastic leukemia; MPN: myeloproliferative neoplasm.

the PRC2-EZH2 axis for survival and TKI resistance.[73,74] Likewise, higher *BMI1* levels at diagnosis correlated with disease progression from CP to BC.[12] while *BMI1* overexpression in CP CD34$^+$ cells increased proliferation and self-renewal,[75] and transformed B-lymphoid progenitors *in vivo*.[76]

DNA methylation-associated gene expression changes

Many studies have examined the role of DNA methylation as a regulator of aberrant GE in CML pathogenesis. In candidate-based approaches, genes involved in cell cycle regulation (*P16*, *P53*, *PLCD1*, *PER3*, *HIC1*), differentiation (*HOXA4*, *DLX4*, *DDIT3*, *SPI1*) proliferation (*CDH13*, *DAPK1*), apoptosis (*BIM*), Wnt regulation (*sFRP1*, *CBY1*), LSC maintenance (*MTSS1*), and cell signaling (*Jun B*, *SOCS2*) were identified as targets of DNA methylation.[67,77]

Recent unbiased genome-wide methylome analyses have solidified the concept of aberrant DNA methylation as a driver of resistance and transformation. The number of differentially methylated regions in CP increased from ~600 to ~6,500 CpG sites in BC.[78] BC was associated with heightened DNA hypermethylation, and to a lesser extent hypomethylation, around promoters of genes involved in stem cell fate, differentiation and leukemia-related functions.[10] Mechanistically, differential DNA methylation patterns in CML have been attributed to underlying *DNMT3A/TET2* mutations, PRC2-dependent epigenetic re-programming, and cytosolic sequestration of Tet2 by BCR-ABL1.[79] Notably, the physiological targeting of DNA hypermethylation using 5-aza-2'-deoxycytidine ameliorated disease phenotypes in a mouse model of CP disease,[80] while low-dose decitabine displayed clinical activity in patients refractory to imatinib,[81] suggesting DNA methylation does indeed contribute to TKI resistance.

Based on the biological insights gleaned so far, it is possible that progression-related DNA methylation signatures may already be evident at diagnosis, particularly in patients presenting with advanced CP.[10] The DNA methylation status of specific target genes might therefore be useful in the timely identification of such patients for more aggressive therapies. Furthermore, given that DNA methylation is a relatively stable epigenetic and biochemical mark, there are practical advantages to developing DNA methylation-based biomarkers rather than transcript-based readouts, especially for the development of robust clinical-grade tests (Figure 3).

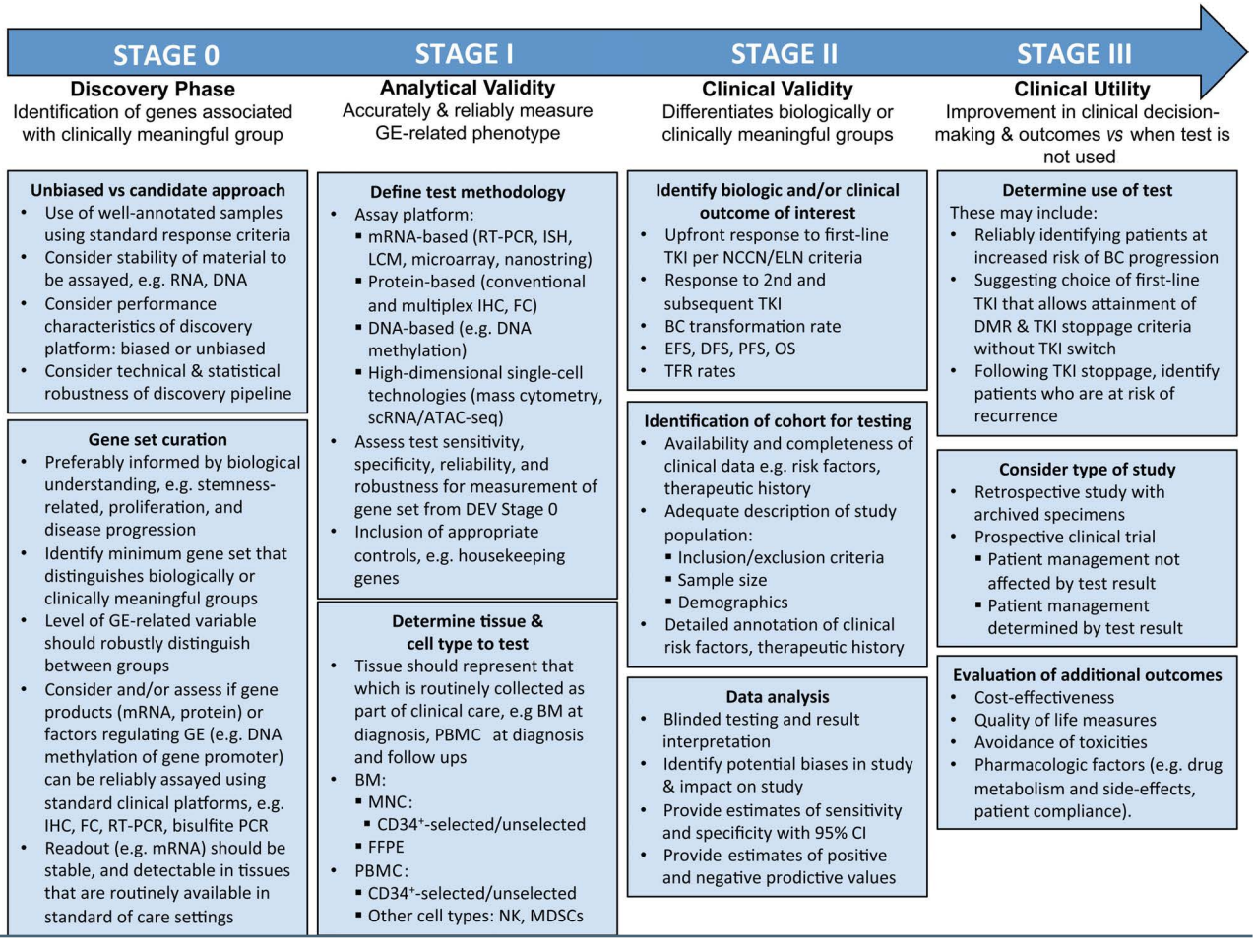

Figure 3. Stages of development of gene expression-based biomarkers. In chronic myeloid leukemia (CML), the development of gene expression-based biomarkers can be divided into three stages following an initial discovery phase. These stages will each determine the analytical validity, clinical validity, and clinical utility of the tests in question. Examples of CML-specific issues or questions that are pertinent to each stage are outlned in boxes under each stage. GE: gene expression; IHC: immunohistochemistry; FC: flow cytometry; RT-PCR: reverse transcriptase polymerase chain reaction; ISH: *in-situ* hybridization; LCM: laser capture microdissection; scRNA-seq: single-cell RNA sequencing; ATAC-seq: assay for transposase-accessible chromatin sequencing; BM: bone marrow; PBMC: peripheral blood mononuclear cell; MNC: mononuclear cells; FFPE: formalin-fixed paraffin-embedded tissues; PB: peripheral blood; NK: natural killer cells; MDSC: myeloid-derived suppressor cells; TKI: tyrosine kinase inhibitor; NCCN: National Comprehensive Cancer Network; ELN: European LeukemiaNet; BC: blast crisis; EFS: event-free survival; DFS: disease-free survival; PFS: progression-free survival; OS: overall survival, TFR: treatment-free remission; 95% CI: 95% confidence interval; DMR: deep molecular response..

Gene expression profiles and mutations: 'seed and soil' revisited

As described above, it will be important to develop CML models that integrate the interaction between genetic and epigenetic factors in driving drug resistance and disease transformation. In this respect, the effects of specific mutations may be cell-context dependent, with differential effects on GE and function depending on the cell type being examined. This is particularly the case for mutations affecting transcription factors, for which cell states, and their attendant chromatin accessibility profile, determine whether the mutated transcription factor has access to its target genes.

To integrate contributions from both the above features, we propose a model in which the cell of origin, with its attendant epigenetic and transcriptional program, determines the ability of specific mutations to contribute to biological and clinical outcomes (Figure 2). This model is a derivative of the 'seed and soil' hypothesis of cancer initiation.[82] The model will be useful for hypothesis testing, and likely explains an important feature of BCR-ABL1 itself. It has been shown in murine models that only when expressed in HSC, but not more committed progenitors, can BCR-ABL1 induce a myeloproliferative disorder. This is likely because BCR-ABL1 is incapable of conferring self-renewal capacity upon committed progenitors, indicating that CML cells rely on BCR-ABL1-independent mechanisms for stemness programs. These findings are in contrast to those for other leukemia fusion genes (e.g., *MLL-ENL*, *MLL-AF9*, *MOZ-TIF2*) which are capable of conferring self-renewal and transform progenitor cells.[83] Relatedly, the model may also explain a naturally occurring phenomenon whereby normal individuals found to carry the *BCR-ABL1* fusion in their peripheral blood mononuclear cells apparently never develop CML.[84] Here, the model would posit that the *BCR-ABL1* fusion is occurring in a long-lived progenitor without self-renewal function.

Analogous to the situation regarding cancer initiation by leukemia fusion genes, mutations devoid of self-renewal function may only confer an increased risk of BC transformation when they occur in a target cell that already possesses physiological self-renewal function. According to this model, mutations in *RUNX1* that are sufficient to induce BC-like disease in mice (Table 3) may be deemed a 'strong' biological seed that can transform many cell types within the hematopoietic hierarchy. Such mutations would be expected to induce disease progression in the majority of patients who harbor such mutations, which is indeed the case.[37] However, a minority of CP patients with *RUNX1* mutations continue to enjoy sustained deep maolecular responses,[57] suggesting the existence of other important factors that modulate *RUNX1* function. Along the same lines, *ASXL1* was recently identified as the most frequently mutated gene at diagnosis in nine patients, the majority (n=6) of whom eventually developed BC, while a minority (n=2) achieved a MMR.[14]

In contrast to the above examples, the prognostic impact of 'weak' seeds is much less clear. In a study by Kim *et al.*, at least four different patterns were observed for *TET2* mutations.[57] One pattern is seen in patients with TKI resistance when both *TET2* and *ABL1* variant allele frequencies increased following TKI therapy, while another is seen when the *TET2* variant allele frequency reduces after TKI treatment in patients with disease progression. In other cases, *TET2* mutations were also detected within Philadelphia chromosome-negative cells, and here, patients showed complex outcomes following TKI therapy, with some achieving MMR and others showing TKI resistance. These observations suggest that the effect of *TET2* mutations are highly contextual.

Challenges ahead but room for optimism

As described above, the discovery of a limited and tractable set of genes that is prognostic across a majority of CML patients has been challenging for clinical, biological, and technical reasons. Nevertheless, there is room for optimism. In the setting of breast cancer, GE panels comprising 21 genes that encompass various aspects of breast cancer biology have been found to be predictive of therapeutic response, and minimized the use of additional therapy without compromising survival.[85] Among liquid tumors, a recent study in acute myeloid leukemia demonstrated that a parsimonious 17-gene GE score, derived from a larger set of stemness-conferring genes, predicts resistance to initial therapy.[86] Interestingly, this score was independent of cytogenetic and mutational risk factors, and suggests that biological factors (e.g., stemness) transcend traditional genetics-based groupings.[87]

Encouragingly in CML, two recent reports suggest that it is possible, using peripheral blood samples taken at diagnosis or 3 months after diagnosis, to predict deep molecular responses and also sustained treatment-free remissions. In the first study, the Adelaide group showed that the rate of decline of *BCR-ABL1* transcripts during first-line TKI therapy (calculated from baseline and 3-month *BCR-ABL1* transcript levels) predicts success of treatment-free remission.[88] The time taken for BCR-ABL1 transcripts to halve was the strongest independent predictor of sustained treatment-free remission: 80% in patients with a halving time of <9.35 days *versus* 4% if the halving time was >21.85 days (P<0.001). In a separate study, Radich *et al.* reported that GE signatures from peripheral blood taken prior to TKI initiation can distinguish individuals who will achieve a deep molecular response (MR4.5) at 5 years from those who will have suboptimal responses.[89] Thus, biological information encoded in GE data can predict very long-term clinical outcomes in CML, and it is therefore conceivable that GE-based data will be able to identify patients in whom TKI therapy can be safely discontinued. More importantly, these early reports suggest that despite the likely existence of diverse resistance mechanisms within the study populations, final common paths, readout either as dynamic measures of *BCR-ABL1* transcript levels, or peripheral blood GE signatures are indeed discoverable.

Stages in developing gene expression-based risk assessment

The stages of developing GE-based tests has been outlined in recent reviews and consensus statements, and comprise at least three phases that assess: analytical validity (reliably measuring the genotype of interest), clinical validity (ability to segregate patients into biologi-

cally and clinically important subsets), and clinical utility (ability to improve clinical decision making).[90,91] In this section, we summarize the pertinent stages and highlight issues of particular relevance to GE-based biomarker development in CML (Figure 3).

Stage 0 is the discovery stage, which is where the field is currently. Here, we highlight three important components, which include the use of technical approaches for unbiased discovery, the simultaneous interrogation of leukemic and non-leukemic clones from the same sample (since both have been shown to be prognostic), and the use of robust statistical and computational pipelines to discover minimal prognostic genes sets. The advent of single-cell-based technologies and their application to well-annotated cohorts will facilitate this step.

In stage I, the minimal gene set has to be converted into a clinical test that accurately and reproducibly measures the GE phenotype. The test platform needs to be robust, as well as sensitive, specific and reliable. The assay should be developed for tissues that are collected as part of routine clinical care. Ideally, any additional processing of material beyond what is routine should be minimized, e.g., $CD34^+$ selection, and should utilize standard procedures available in clinical laboratories, such as flow cytometry and bone marrow immunohistochemistry. An example would be detecting GE signatures of interest by a panel of antibodies for use in flow cytometry or immunohistochemistry applications. It is preferable that the samples used for analytical validation are from well-characterized patients representative of 'real-world' settings and, ideally, validated in at least one independent cohort. Sample size and power calculations should be determined prior to starting the study, and analytic sensitivity and specificity for the test should be available at the end of the study. At the end of stage I, a locked-down test should be evaluated in stage II, that of clinical validation.

In stage II, the locked-down test will be evaluated for its ability to differentiate between clinically meaningful outcomes in modern CML practice. The samples to be tested should be obtained from well-annotated cohorts representative of the broader population, and the test conducted on tissues in a blinded manner with respect to testing and result reporting. Ideal populations include patients who have been treated uniformly in clinical trials. At the end of this stage, the ability of the test to predict clinical outcome should be available as a test score, with clearly defined positive and negative predictive values.

The final stage, stage III, will be the determination of clinical utility. This stage would entail the use of the GE-based test to improve clinical decision-making, and would require the study to demonstrate that meaningful outcomes are improved when the test is used compared to when the test is not used. Besides clinical outcomes such as improved progression-free survival and overall survival, additional measures such as cost-effectiveness, avoidance of toxicities, quality of life and psychological parameters should also be assessed. Such studies may also incorporate the contribution of pharmacological factors (e.g., drug metabolism and side effects, patient compliance) to overall outcomes. Given the relative rarity of CML, it is envisaged that this will be a multicenter international study.

Conclusion

Genetic and epigenetic events contribute to the emergence of BCR-ABL1-independent clones that result in clinical TKI resistance and, if unopposed, BC transformation. Long-term TKI responses, including successful TKI stoppage, can be predicted by slower declines in *BCR-ABL1* transcript levels during first-line TKI therapy,[88] suggesting that genetic and epigenetic factors contributing to TKI resistance are present at diagnosis. Recent studies describe a convergent GE signature common to the majority of BC progenitors.[10] Elements of this common or core transcriptome can be detected in $CD34^+$ cells from CP patients at risk of TKI resistance or early transformation,[11,13] and specific mutations have been shown to contribute additional nuances to the core transcriptome.[49] These observations are consistent with a 'seed and soil' model that may be helpful for hypothesis generation (Figure 2). Emerging technologies, particularly multimodal single-cell-based approaches, will facilitate the discovery of genetic and epigenetic biomarkers at presentation. This initial discovery phase has to be followed by the translation of GE-based information into validated analytical tests, and subsequently, the determination of clinical validity and utility. This process will be a multi-year, multi-institution international effort akin to that for the development of a genetic-based risk assessment.[85,90,91] The integration of both gene mutation- and gene expression-based biomarkers into the care of CML patients will be an important step to achieving the ultimate goal of CML research: the cure of the majority of our patients.

Disclosures

DK is a member of the advisory boards of Novartis, Pfizer, Paladin, and has received honoraria from Novartis, Pfizer and Paladin, as well as research funding from Novartis, Bristol-Myers Squibb, Pfizer, and Paladin. TH is a member of a Novartis advisory board and receives research support from Novartis and Bristol-Myers Squibb. SB is a member of the advisory boards of Qiagen, Novartis, and Cepheid and has received honoraria from Qiagen, Novartis, Bristol-Myers Squibb, and Cepheid, as well as research support from Novartis and Cepheid.

Contributions

VK and STO conceived the topic for review, and wrote the first draft of the manuscript. DDHK, TPH, and SB contributed by the addition of new sections and critical discussions throughout the writing of the review.

Funding

VK and STO are supported by the National Medical Research Council Singapore (MOH-CSASI18may-0002, MOH-CIRG20nov-0003, NMRC/CIRG/1468/2017).

References

1. Rowley JD. Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. Nature. 1973;243(5405):290-293.
2. Perrotti D, Jamieson C, Goldman J, Skorski T. Chronic myeloid leukemia: mechanisms of blastic transformation. J Clin Invest. 2010;120(7):2254-2264.
3. Druker BJ. Translation of the Philadelphia chromosome into therapy for CML. Blood. 2008;112(13):4808-4817.
4. Holyoake TL, Vetrie D. The chronic myeloid leukemia stem cell: stemming the tide of persistence. Blood. 2017;129(12): 1595-1606.
5. Hochhaus A, Baccarani M, Silver RT, et al. European LeukemiaNet 2020 recommendations for treating chronic myeloid leukemia. Leukemia. 2020;34(4):966-984.
6. Mahon FX, Rea D, Guilhot J, et al. Discontinuation of imatinib in patients with chronic myeloid leukaemia who have maintained complete molecular remission for at least 2 years: the prospective, multicentre Stop Imatinib (STIM) trial. Lancet Oncol. 2010;11(11):1029-1035.
7. Ross DM, Hughes TP. Treatment-free remission in patients with chronic myeloid leukaemia. Nat Rev Clin Oncol. 2020;17 (8):493-503.
8. Branford S, Fletcher L, Cross NC, et al. Desirable performance characteristics for BCR-ABL measurement on an international reporting scale to allow consistent interpretation of individual patient response and comparison of response rates between clinical trials. Blood. 2008;112(8):3330-3338.
9. Malhotra H, Radich J, Garcia-Gonzalez P. Meeting the needs of CML patients in resource-poor countries. Hematology Am Soc Hematol Educ Program. 2019;2019 (1):433-442.
10. Ko TK, Javed A, Lee KL, et al. An integrative model of pathway convergence in genetically heterogeneous blast crisis chronic myeloid leukemia. Blood. 2020;135 (26):2337-2353.
11. McWeeney SK, Pemberton LC, Loriaux MM, et al. A gene expression signature of CD34+ cells to predict major cytogenetic response in chronic-phase chronic myeloid leukemia patients treated with imatinib. Blood. 2010;115(2):315-325.
12. Mohty M, Yong AS, Szydlo RM, Apperley JF, Melo JV. The polycomb group BMI1 gene is a molecular marker for predicting prognosis of chronic myeloid leukemia. Blood. 2007;110(1):380-383.
13. Yong AS, Szydlo RM, Goldman JM, Apperley JF, Melo JV. Molecular profiling of CD34+ cells identifies low expression of CD7, along with high expression of proteinase 3 or elastase, as predictors of longer survival in patients with CML. Blood. 2006;107(1):205-212.
14. Branford S, Wang P, Yeung DT, et al. Integrative genomic analysis reveals cancer-associated mutations at diagnosis of CML in patients with high-risk disease. Blood. 2018;132(9):948-961.
15. Malone ER, Oliva M, Sabatini PJB, Stockley TL, Siu LL. Molecular profiling for precision cancer therapies. Genome Med. 2020;12 (1):8.
16. Kaneta Y, Kagami Y, Katagiri T, et al. Prediction of sensitivity to STI571 among chronic myeloid leukemia patients by genome-wide cDNA microarray analysis. Jpn J Cancer Res. 2002;93(8):849-856.
17. McLean LA, Gathmann I, Capdeville R, Polymeropoulos MH, Dressman M. Pharmacogenomic analysis of cytogenetic response in chronic myeloid leukemia patients treated with imatinib. Clin Cancer Res. 2004;10(1 Pt 1):155-165.
18. de Lavallade H, Finetti P, Carbuccia N, et al. A gene expression signature of primary resistance to imatinib in chronic myeloid leukemia. Leuk Res. 2010;34(2):254-257.
19. Zhang WW, Cortes JE, Yao H, et al. Predictors of primary imatinib resistance in chronic myelogenous leukemia are distinct from those in secondary imatinib resistance. J Clin Oncol. 2009;27(22):3642-3649.
20. Kok CH, Yeung DT, Lu L, et al. Gene expression signature that predicts early molecular response failure in chronic-phase CML patients on frontline imatinib. Blood Adv. 2019;3(10):1610-1621.
21. Marin D, Ibrahim AR, Lucas C, et al. Assessment of BCR-ABL1 transcript levels at 3 months is the only requirement for predicting outcome for patients with chronic myeloid leukemia treated with tyrosine kinase inhibitors. J Clin Oncol. 2012;30(3): 232-238.
22. Hughes TP, Saglio G, Kantarjian HM, et al. Early molecular response predicts outcomes in patients with chronic myeloid leukemia in chronic phase treated with frontline nilotinib or imatinib. Blood. 2014;123(9):1353-1360.
23. Frank O, Brors B, Fabarius A, et al. Gene expression signature of primary imatinib-resistant chronic myeloid leukemia patients. Leukemia. 2006;20(8):1400-1407.
24. Villuendas R, Steegmann JL, Pollan M, et al. Identification of genes involved in imatinib resistance in CML: a gene-expression profiling approach. Leukemia. 2006;20(6): 1047-1054.
25. Radich JP, Dai H, Mao M, et al. Gene expression changes associated with progression and response in chronic myeloid leukemia. Proc Natl Acad Sci U S A. 2006;103(8):2794-2799.
26. Lim S, Saw TY, Zhang M, et al. Targeting of the MNK-eIF4E axis in blast crisis chronic myeloid leukemia inhibits leukemia stem cell function. Proc Natl Acad Sci U S A. 2013;110(25):E2298-2307.
27. Crossman LC, Mori M, Hsieh YC, et al. In chronic myeloid leukemia white cells from cytogenetic responders and non-responders to imatinib have very similar gene expression signatures. Haematologica. 2005;90(4): 459-464.
28. Burguillo FJ, Martin J, Barrera I, Bardsley WG. Meta-analysis of microarray data: the case of imatinib resistance in chronic myelogenous leukemia. Comput Biol Chem. 2010;34(3):184-192.
29. Jamieson CH, Ailles LE, Dylla SJ, et al. Granulocyte-macrophage progenitors as candidate leukemic stem cells in blast-crisis CML. N Engl J Med. 2004;351(7):657-667.
30. Lim B, Lin Y, Navin N. Advancing cancer research and medicine with single-cell genomics. Cancer Cell. 2020;37(4):456-470.
31. Warfvinge R, Geironson L, Sommarin MNE, et al. Single-cell molecular analysis defines therapy response and immunophenotype of stem cell subpopulations in CML. Blood. 2017;129(17):2384-2394.
32. Giustacchini A, Thongjuea S, Barkas N, et al. Single-cell transcriptomics uncovers distinct molecular signatures of stem cells in chronic myeloid leukemia. Nat Med. 2017;23(6):692-702.
33. Inselmann S, Wang Y, Saussele S, et al. Development, function, and clinical significance of plasmacytoid dendritic cells in chronic myeloid leukemia. Cancer Res. 2018;78(21):6223-6234.
34. Sinnakannu JR, Lee KL, Cheng S, et al. SRSF1 mediates cytokine-induced impaired imatinib sensitivity in chronic myeloid leukemia. Leukemia. 2020;34(7):1787-1798.
35. Oehler VG, Yeung KY, Choi YE, et al. The derivation of diagnostic markers of chronic myeloid leukemia progression from microarray data. Blood. 2009;114(15):3292-3298.
36. Zheng C, Li L, Haak M, et al. Gene expression profiling of CD34+ cells identifies a molecular signature of chronic myeloid leukemia blast crisis. Leukemia. 2006;20(6): 1028-1034.
37. Branford S, Kim DDH, Apperley JF, et al. Laying the foundation for genomically-based risk assessment in chronic myeloid leukemia. Leukemia. 2019;33(8):1835-1850.
38. Zhao LJ, Wang YY, Li G, et al. Functional features of RUNX1 mutants in acute transformation of chronic myeloid leukemia and their contribution to inducing murine full-blown leukemia. Blood. 2012;119(12):2873-2882.
39. Joshi I, Yoshida T, Jena N, et al. Loss of Ikaros DNA-binding function confers integrin-dependent survival on pre-B cells and progression to acute lymphoblastic leukemia. Nat Immunol. 2014;15(3):294-304.
40. Balasubramani A, Larjo A, Bassein JA, et al. Cancer-associated ASXL1 mutations may act as gain-of-function mutations of the ASXL1-BAP1 complex. Nat Commun. 2015;6:7307.
41. Pagan JK, Arnold J, Hanchard KJ, et al. A novel corepressor, BCoR-L1, represses transcription through an interaction with CtBP. J Biol Chem. 2007;282(20):15248-15257.
42. Kazenwadel J, Secker GA, Liu YJ, et al. Loss-of-function germline GATA2 mutations in patients with MDS/AML or MonoMAC syndrome and primary lymphedema reveal a key role for GATA2 in the lymphatic vasculature. Blood. 2012;119(5):1283-1291.
43. Zhang SJ, Shi JY, Li JY. GATA-2 L359V mutation is exclusively associated with CML progression but not other hematological malignancies and GATA-2 P250A is a novel single nucleotide polymorphism. Leuk Res. 2009;33(8):1141-1143.
44. Pronier E, Almire C, Mokrani H, et al. Inhibition of TET2-mediated conversion of 5-methylcytosine to 5-hydroxymethylcytosine disturbs erythroid and granulomonocytic differentiation of human hematopoietic progenitors. Blood. 2011; 118(9):2551-2555.
45. Moran-Crusio K, Reavie L, Shih A, et al. Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. Cancer Cell. 2011;20(1):11-24.
46. Mayle A, Yang L, Rodriguez B, et al. Dnmt3a loss predisposes murine hematopoietic stem cells to malignant transformation. Blood. 2015;125(4):629-638.
47. Hervouet E, Peixoto P, Delage-Mourroux R, Boyer-Guittaut M, Cartron PF. Specific or not specific recruitment of DNMTs for DNA methylation, an epigenetic dilemma. Clin Epigenetics. 2018;10:17.
48. Huang Y-H, Tovy A, Sundaramurthy V, et al. Nearly a third of clonal hematopoiesis-

associated DNMT3A mutations reduce protein stability and may be associated with poorer prognosis. Blood. 2018;132 (Suppl 1):1315.
49. Adnan Awad S, Dufva O, Ianevski A, et al. RUNX1 mutations in blast-phase chronic myeloid leukemia associate with distinct phenotypes, transcriptional profiles, and drug responses. Leukemia. 2021;35(4): 1087-1099.
50. Warsch W, Kollmann K, Eckelhart E, et al. High STAT5 levels mediate imatinib resistance and indicate disease progression in chronic myeloid leukemia. Blood. 2011;117 (12):3409-3420.
51. Beer PA, Knapp DJ, Miller PH, et al. Disruption of IKAROS activity in primitive chronic-phase CML cells mimics myeloid disease progression. Blood. 2015;125(3): 504-515.
52. Thomson DW, Shahrin NH, Wang PPS, et al. Aberrant RAG-mediated recombination contributes to multiple structural rearrangements in lymphoid blast crisis of chronic myeloid leukemia. Leukemia. 2020;34(8): 2051-2063.
53. Giotopoulos G, van der Weyden L, Osaki H, et al. A novel mouse model identifies cooperating mutations and therapeutic targets critical for chronic myeloid leukemia progression. J Exp Med. 2015;212(10):1551-1569.
54. Yang H, Kurtenbach S, Guo Y, et al. Gain of function of ASXL1 truncating protein in the pathogenesis of myeloid malignancies. Blood. 2018;131(3):328-341.
55. Jaiswal S, Ebert BL. Clonal hematopoiesis in human aging and disease. Science. 2019; 366(6465):eaan4673.
56. Schmidt M, Rinke J, Schafer V, et al. Molecular-defined clonal evolution in patients with chronic myeloid leukemia independent of the BCR-ABL status. Leukemia. 2014;28(12):2292-2299.
57. Kim T, Tyndel MS, Kim HJ, et al. Spectrum of somatic mutation dynamics in chronic myeloid leukemia following tyrosine kinase inhibitor therapy. Blood. 2017;129 (1):38-47.
58. Jeong M, Sun D, Luo M, et al. Large conserved domains of low DNA methylation maintained by Dnmt3a. Nat Genet. 2014;46(1):17-23.
59. Abdel-Wahab O, Gao J, Adli M, et al. Deletion of Asxl1 results in myelodysplasia and severe developmental defects in vivo. J Exp Med. 2013;210(12):2641-2659.
60. Sano S, Oshima K, Wang Y, et al. CRISPR-mediated gene editing to assess the roles of Tet2 and Dnmt3a in clonal hematopoiesis and cardiovascular disease. Circ Res. 2018;123(3):335-341.
61. Sano S, Oshima K, Wang Y, et al. Tet2-mediated clonal hematopoiesis accelerates heart failure through a mechanism involving the IL-1beta/NLRP3 inflammasome. J Am Coll Cardiol. 2018;71(8):875-886.
62. Arends CM, Galan-Sousa J, Hoyer K, et al. Hematopoietic lineage distribution and evolutionary dynamics of clonal hematopoiesis. Leukemia. 2018;32(9):1908-1919.
63. Welner RS, Amabile G, Bararia D, et al. Treatment of chronic myelogenous leukemia by blocking cytokine alterations found in normal stem and progenitor cells. Cancer Cell. 2015;27(5):671-681.
64. Hughes A, Yong ASM. Immune effector recovery in chronic myeloid leukemia and treatment-free remission. Front Immunol. 2017;8:469.
65. Hughes A, Clarson J, Tang C, et al. CML patients with deep molecular responses to TKI have restored immune effectors and decreased PD-1 and immune suppressors. Blood. 2017;129(9):1166-1176.
66. Yong AS, Keyvanfar K, Hensel N, et al. Primitive quiescent CD34+ cells in chronic myeloid leukemia are targeted by in vitro expanded natural killer cells, which are functionally enhanced by bortezomib. Blood. 2009;113(4):875-882.
67. Koschmieder S, Vetrie D. Epigenetic dysregulation in chronic myeloid leukaemia: a myriad of mechanisms and therapeutic options. Semin Cancer Biol. 2018;51:180-197.
68. Sparmann A, van Lohuizen M. Polycomb silencers control cell fate, development and cancer. Nat Rev Cancer. 2006;6(11):846-856.
69. Katoh M. Functional and cancer genomics of ASXL family members. Br J Cancer. 2013;109(2):299-306.
70. Wong SJ, Gearhart MD, Taylor AB, et al. KDM2B recruitment of the polycomb group complex, PRC1.1, requires cooperation between PCGF1 and BCORL1. Structure. 2016;24(10):1795-1801.
71. Yu M, Mazor T, Huang H, et al. Direct recruitment of polycomb repressive complex 1 to chromatin by core binding transcription factors. Mol Cell. 2012;45(3):330-343.
72. Oravecz A, Apostolov A, Polak K, et al. Ikaros mediates gene silencing in T cells through Polycomb repressive complex 2. Nat Commun. 2015;6:8823.
73. Xie H, Peng C, Huang J, et al. Chronic myelogenous leukemia-initiating cells require polycomb group protein EZH2. Cancer Discov. 2016;6(11):1237-1247.
74. Scott MT, Korfi K, Saffrey P, et al. Epigenetic reprogramming sensitizes CML stem cells to combined EZH2 and tyrosine kinase inhibition. Cancer Discov. 2016;6(11):1248-1257.
75. Rizo A, Horton SJ, Olthof S, et al. BMI1 collaborates with BCR-ABL in leukemic transformation of human CD34+ cells. Blood. 2010;116(22):4621-4630.
76. Sengupta A, Ficker AM, Dunn SK, Madhu M, Cancelas JA. Bmi1 reprograms CML B-lymphoid progenitors to become B-ALL-initiating cells. Blood. 2012;119(2):494-502.
77. Behzad MM, Shahrabi S, Jaseb K, et al. Aberrant DNA methylation in chronic myeloid leukemia: cell fate control, prognosis, and therapeutic response. Biochem Genet. 2018;56(3):149-175.
78. Heller G, Topakian T, Altenberger C, et al. Next-generation sequencing identifies major DNA methylation changes during progression of Ph+ chronic myeloid leukemia. Leukemia. 2016;30(9):1861-1868.
79. Mancini M, Veljkovic N, Leo E, et al. Cytoplasmatic compartmentalization by Bcr-Abl promotes TET2 loss-of-function in chronic myeloid leukemia. J Cell Biochem. 2012;113(8):2765-2774.
80. Amabile G, Di Ruscio A, Muller F, et al. Dissecting the role of aberrant DNA methylation in human leukaemia. Nat Commun. 2015;6:7091.
81. Issa JP, Gharibyan V, Cortes J, et al. Phase II study of low-dose decitabine in patients with chronic myelogenous leukemia resistant to imatinib mesylate. J Clin Oncol. 2005;23(17):3948-3956.
82. Visvader JE. Cells of origin in cancer. Nature. 2011;469(7330):314-322.
83. Huntly BJ, Shigematsu H, Deguchi K, et al. MOZ-TIF2, but not BCR-ABL, confers properties of leukemic stem cells to committed murine hematopoietic progenitors. Cancer Cell. 2004;6(6):587-596.
84. Bose S, Deininger M, Gora-Tybor J, Goldman JM, Melo JV. 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;92(9):3362-3367.
85. Sparano JA, Gray RJ, Makower DF, et al. Adjuvant chemotherapy guided by a 21-gene expression assay in breast cancer. N Engl J Med. 2018;379(2):111-121.
86. Ng SW, Mitchell A, Kennedy JA, et al. A 17-gene stemness score for rapid determination of risk in acute leukaemia. Nature. 2016;540(7633):433-437.
87. Bill M, Nicolet D, Kohlschmidt J, et al. Mutations associated with a 17-gene leukemia stem cell score and the score's prognostic relevance in the context of the European LeukemiaNet classification of acute myeloid leukemia. Haematologica. 2020;105(3):721-729.
88. Shanmuganathan N, Pagani IS, Ross DM, et al. Early BCR-ABL1 kinetics are predictive of subsequent achievement of treatment-free remission in chronic myeloid leukemia. Blood. 2021;137(9):1196-1207.
89. Radich JP, Larson R, Kantarjian H, et al. Gene expression signature predicts deep molecular response (DMR) in chronic myeloid leukemia (CML): an exploratory biomarker analysis from ENESTnd [Abstract]. Blood. 2019;34;(Suppl_1):665.
90. Kwa M, Makris A, Esteva FJ. Clinical utility of gene-expression signatures in early stage breast cancer. Nat Rev Clin Oncol. 2017;14(10):595-610.
91. Teutsch SM, Bradley LA, Palomaki GE, et al. The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Initiative: methods of the EGAPP Working Group. Genet Med. 2009;11(1):3-14.