Analysis of BCR–ABL1 tyrosine kinase domain mutational spectra in primitive chronic myeloid leukemia cells suggests a unique mutator phenotype

Leukemia (2010) **24,** 1817–1821; doi:10.1038/leu.2010.179;
Published online 26 August 2010

Chronic myeloid leukemia (CML) is a clonal hematopoietic stem cell disorder characterized by a *BCR–ABL1* fusion gene, usually resulting from a reciprocal t(9;22) translocation.[1] The identification of this abnormality led to the development of therapies that selectively inhibit the activity of the BCR–ABL1 tyrosine kinase (TK) oncoprotein.[2] Imatinib mesylate (IM), the first such inhibitor, has shown considerable effectiveness in chronic phase CML (CP-CML).[2] Nevertheless, acquired and intrinsic

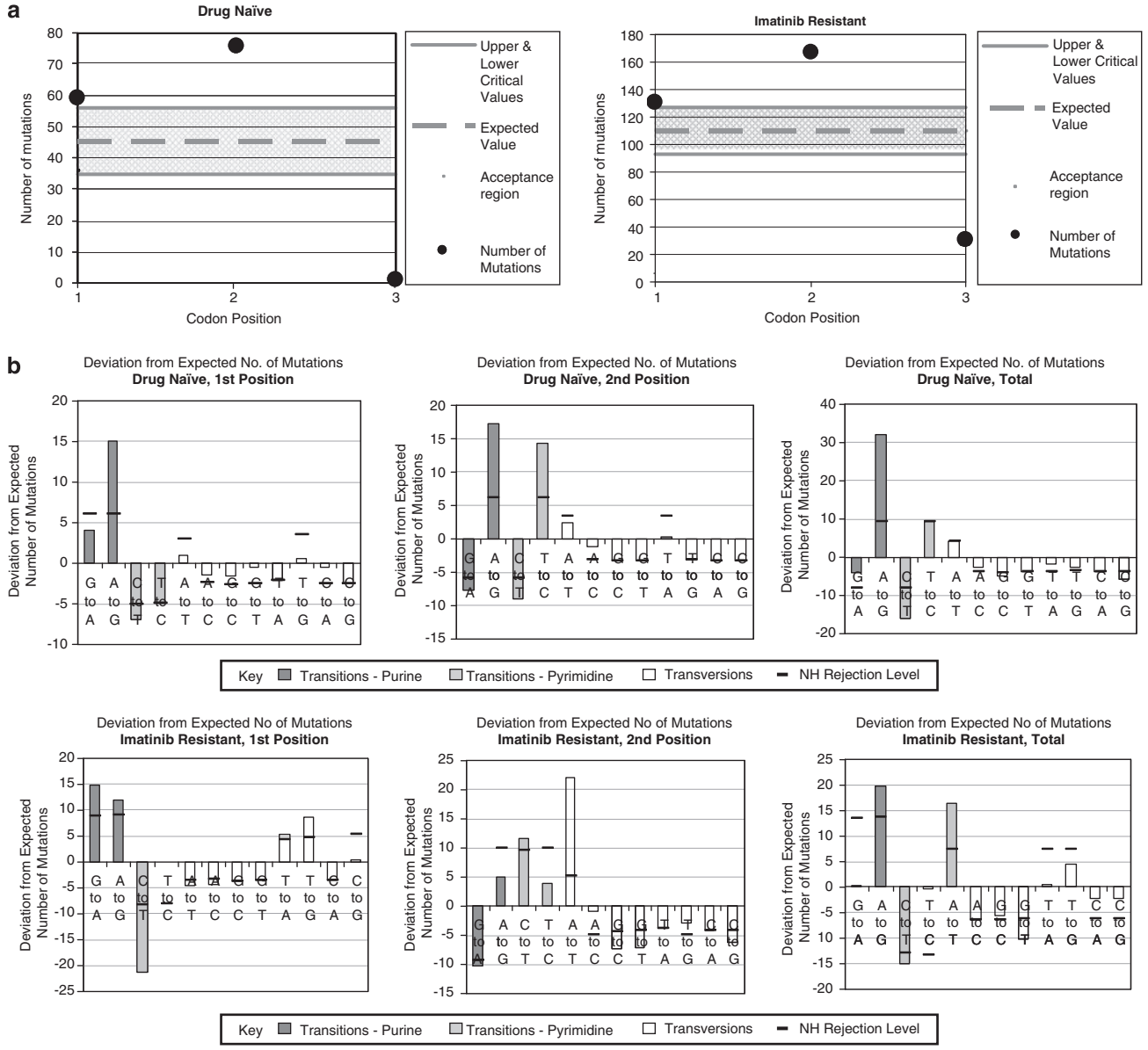

Figure 1 Distribution of *BCR–ABL1* TK domain mutations across codon positions in cells from IM-naïve and IM-resistant CML patients. (**a**) Assuming H_0, the central dashed line shows the expected number of mutations at all three codon positions, the uniform lines show the upper and lower critical values of the test outside which H_0 is rejected, with the shaded area representing the acceptance region. The *P*-values corresponding to codon positions 1, 2 and 3 for the IM-naïve cell data are 0.0191, $1.1\text{x}10^{-7}$ and 1.5×10^{-22}, respectively, and for the IM-resistant cell data are 0.0158, 1.1×10^{-11} and 4.2×10^{-24}, respectively. (**b**) Comparative deviation of each mutation type from the expected frequency derived from unselected regions of the human genome. Transitions and transversions are shaded and the rejection levels of H_0 beyond which the observed numbers of mutations are deemed significant are represented by black dashed lines. Two-tailed tests were performed using a 5% cutoff to reject H_0, in which n_{obs} was significantly different from n_{exp}. Critical test values were generated by a short program in Mathematica (Wolfram Research, http://www.wolfram.com/products/mathematica/index.html). The modified conditional rebalancing/reweighting *P*-value[9] was employed in the calculation of critical values to correct for asymmetry in the binomial distribution.

drug resistance remains a significant clinical problem in CML patients, with amino acid mutations in the *BCR–ABL1* TK domain accounting for 60–90% of relapses.[3] Although some mutations result in amino acid changes conferring IM-resistant phenotypes, others have no significant impact on drug efficacy, but may serve as independent prognostic markers of a poor clinical outcome.[3] Primitive ($CD34^+$) CP-CML cells are intrinsically insensitive to IM and genetically unstable.[4] They persist in IM-treated patients and constitute a reservoir from which IM-resistant mutant clones can be generated. *BCR–ABL1* TK domain mutations are detected in $CD34^+$ CP-CML cells before IM treatment, however, the mechanism by which they arise is unknown.[4]

In this analysis, we studied sequences from 15 IM-naïve and 316 IM-resistant CP-CML patients and identified a total of 460 somatic point mutations in the *BCR–ABL1* TK domain in their leukemic cells. These mutations were non-randomly distributed across the three different codon positions, with very few mutations at codon position 3 (1 of 136 mutations, 0.7%, in IM-naïve patients and 31 of 329, 9%, in IM-resistant patients, Figure 1a). The remaining mutations were not significantly biased in their distribution over codon positions 1 and 2 in IM-naïve patient cells, whereas in IM-resistant cells, a significant positive bias was observed at codon position 2 ($P=0.042$).

On the other hand, *BCR–ABL1* TK domain sequences from IM-naïve patient cells had significantly more transitions relative to unselected regions of the human genome across all codon positions (Figure1b and Table 1). We also noted a significant positive A-to-G mutational bias (A>G, 1.12×10^{-10}) and an under-representation of C-to-T mutations (C>T, 5.34×10^{-5}), when these mutations were compared with unselected regions of the human genome.[5] In addition, we observed a T-to-C mutational hot spot (T>C, 1.23×10^{-4}) at codon position 2. IM-naive cells exhibited a transitional bias at position 2 and overall (1.99×10^{-4}, 6.19×10^{-5}, respectively; at position 1, *P*-value = 0.052 as compared with unselected regions of the human genome, Figures 1b and 2a).[5,6]

We observed a similar pattern of *BCR–ABL1* TK codon position mutations in sequences derived from IM-resistant cells at position 1 and overall. These included a significant A-to-G bias (A>G, $P=0.011$ and 0.006, respectively) and a significant under-representation of C-to-T transitions (C>T, 1.31×10^{-9} and 0.030, respectively, Figure 1b). It is interesting that in IM-resistant cells, A-to-T (A>T) transversions were over-represented at position 2 and overall (1.41×10^{-11} and 3.77×10^{-5}, respectively, Figures 1b and 2b). These cells also had a profound transversional bias (1.45×10^{-4}) at codon position 3, suggesting that positive selection occurs at a position where transversions are typically non-synonymous.[6]

It is interesting that a skewed transition:transversion ratio and transition and transversion frequency at TK codon positions 1 and 2 were observed in IM-naïve and IM-selected cells (Figures 3a and b). The mutations were significantly unevenly distributed across these 2 codon positions. In the TK domain of IM-resistant cells the frequency of C-to-T mutations (C>T, 3.82×10^{-11}) indicated a bias affecting codon position 2 over 1. A-to-T transversions occurred significantly more frequently at position 2 as compared with 1 (A>T, 7.45×10^{-9}). This was also the case for T-to-C transitions derived from TK sequences in IM-naïve cells (T>C, 1.13×10^{-4}). Figure 3c shows higher frequencies of these different mutations in IM-naïve and IM-resistant cells as compared with unselected regions of the human genome,

Table 1 Biases in the frequency of each type of mutation in IM-naïve and -resistant cells compared with unselected regions of the human genome (both for codon positions 1 and 2, and overall across all three codons)

	IM-naïve bias					*IM-resistant bias*[a]				
	First position		*Second position*		*Overall bias*	*First position*		*Second position*		*Overall bias*
Transitions										
G to A (Pu)	None		−'ve	(0.0154)	None	+'ve	(0.0018)	−'ve	(0.0342)	None
A to G (Pu)	+'ve	(6.2×10^{-6})	+'ve	(4.2×10^{-6})	+'ve	+'ve	(0.0119)	None		+'ve
C to T (Py)	−'ve	(0.0115)	−'ve	(0.0046)	−'ve	−'ve	(1.4×10^{-9})	+'ve	(0.0282)	−'ve
T to C (Py)	None		+'ve	(0.0001)	None	None		None		None
Transversions										
A to T	None		None		None	−'ve	(0.0359)	+'ve	(1.4×10^{-11})	+'ve
A to C	None		None		None	−'ve	(0.0498)	None		None
G to C	None		None		−'ve	None		−'ve	(0.001)	None
G to T	None		None		None	None		−'ve	(0.0011)	−'ve
T to A	None		None		None	+'ve	(0.0354)	None		None
T to G	None		None		None	+'ve	(0.0023)	None		None
C to A	None		None		None	None		None		None
C to G	None		None		−'ve	None		−'ve	(0.0197)	None

Abbreviation: IM, imatinib mesylate.
None indicates that the observed number of mutations are comparable to unselected regions, whereas +'ve indicates an over-representation of the particular mutation type and −'ve indicates an under-representation. Pu indicates a purine transition and Py indicates a pyrimidine transition. The corresponding *P*-values are indicated.
[a]The clinical classification of IM-responders are patients who achieved complete hematologic remission within 3 months, a major cytogenetic remission within 12 months and a complete cytogenetic remission within 18 months. Conversely, IM-resistant patients were patients who did not achieve these response criteria or had evidence of loss of a complete hematologic remission or a complete cytogenetic remission.

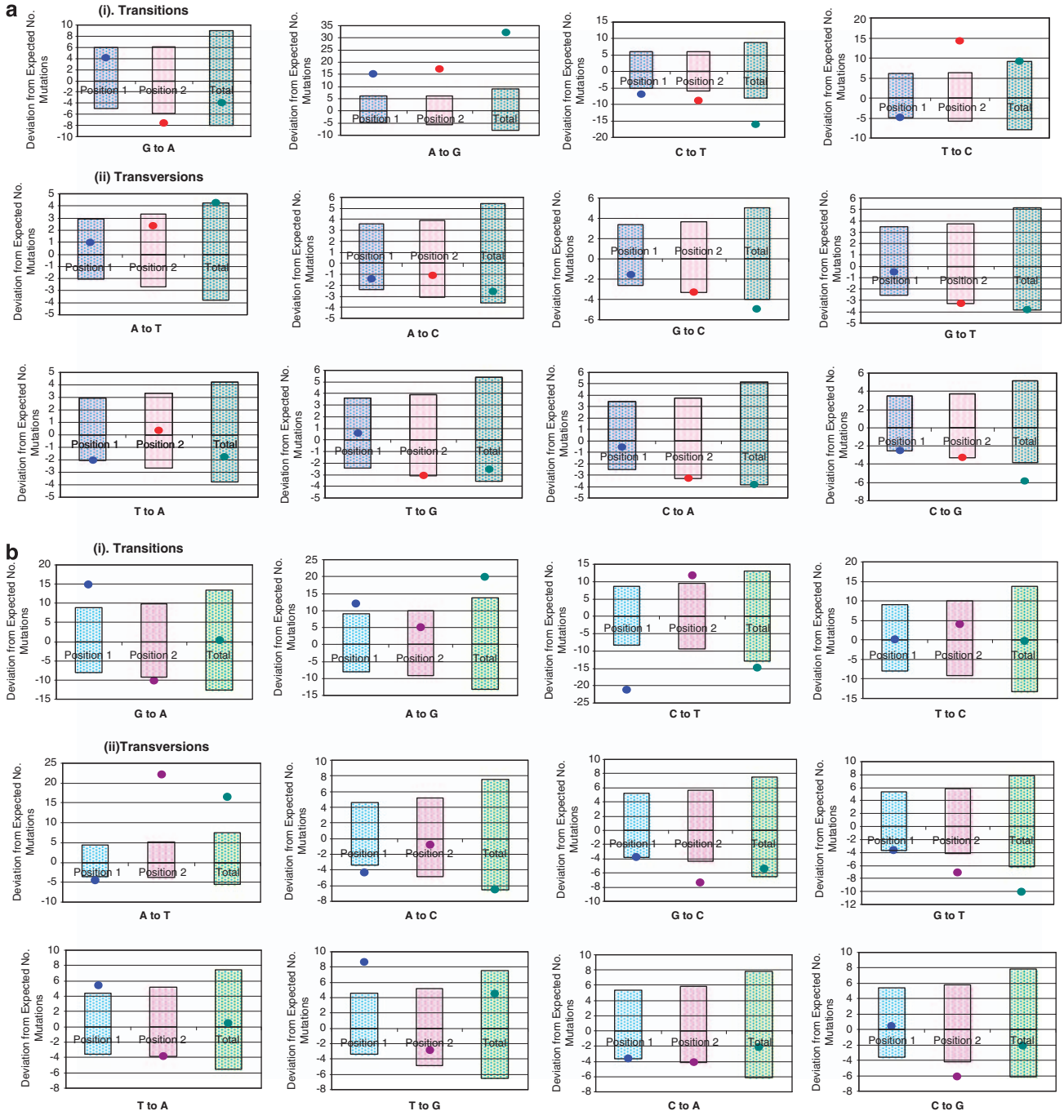

Figure 2 Derivation of mutation frequencies according to mutation type from the number predicted in unselected regions of the human genome. For each codon position from IM-naïve (**a**) and IM-resistant patients (**b**), acceptance regions (shaded) are determined by the null hypothesis indicating that the numbers of mutations are expected in the frequencies found in the unselected regions of the human genome. The large dots represent the observed number of mutations and are given for codon positions 1 and 2, and overall.

providing further evidence that the TK domain mutations in CP-CML have a distinct mutational profile.[5]

Primitive CML cells are characterized by an innate resistance to tyrosine kinase inhibitors (TKIs) and the activity of an unknown mutator that destabilizes their genome and generates somatic point mutations.[4] Our results provide new evidence for the activity of a distinct mutator process in CP-CML. The signature of the mutator comprises an A-to-G transitional bias, A-to-G positional bias with hot spot mutations at codon positions 1 and 2, T-to-C mutations at position 2 and a near lack of position 3 mutations. The clinically observed M244V and D276G mutations result from A-to-G transitions and the F359L mutation arises from T-to-C transitions, both of which would be predicted by the activity of the CML-CP mutator. However, the most clinically important T315I mutation that confers resistance to most currently available TKIs,[7] is generated by a C-to-T transition, suggesting a high mutational rate generating mutational escape around the principal mutator

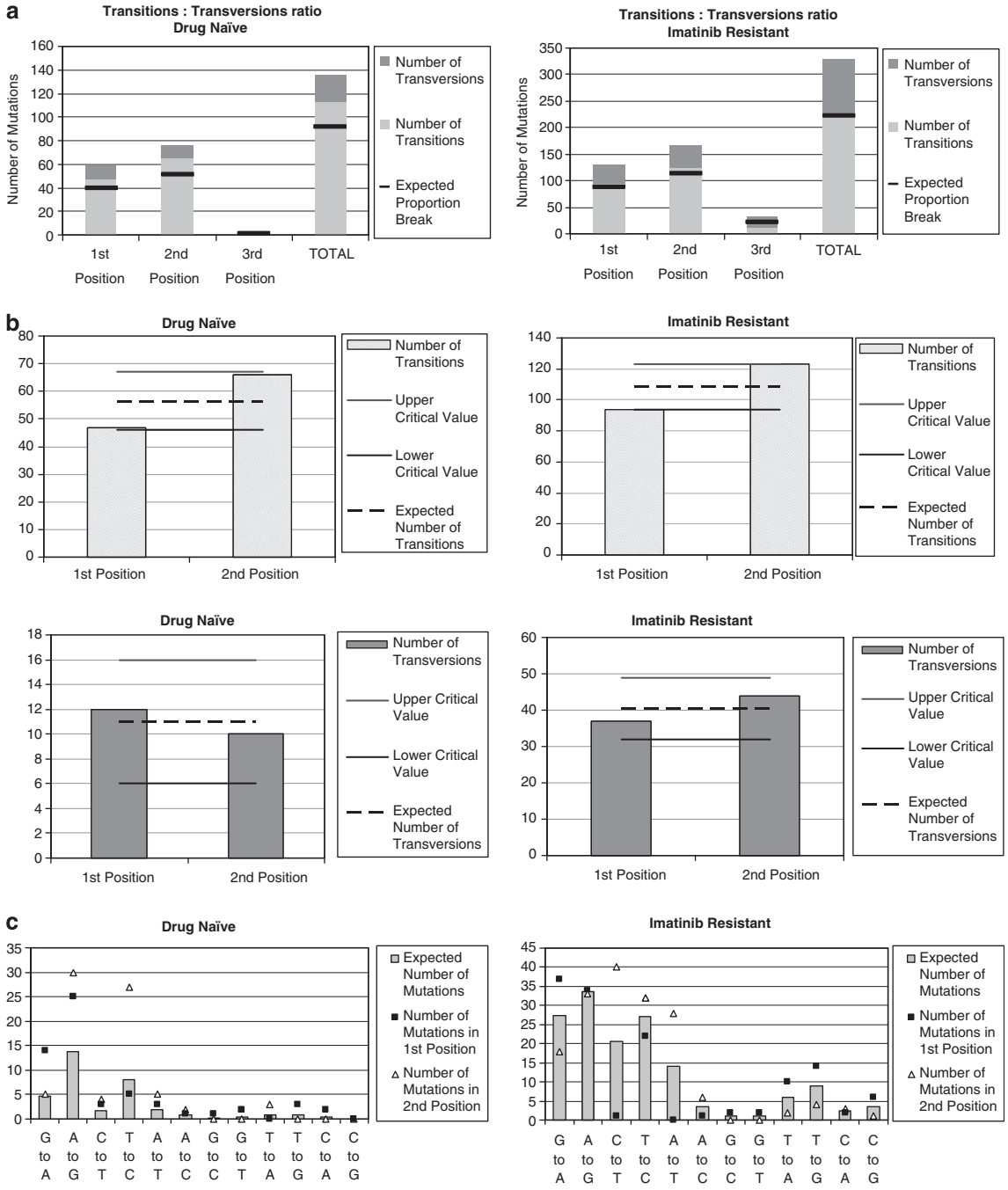

Figure 3 Comparison of the observed transitions:transversions ratio in the TK domain of *BCR–ABL1* with unselected regions of the human genome. (**a**) The total number of mutations at each codon position is subdivided into transversions (dark grey) and transitions (light grey). The black line shows the expected 67.5:32.5 proportion break observed in the unselected regions of the human genome. The *P*-values corresponding to codon positions 1, 2, 3 and 'in total' for the IM-naïve cell data are 0.0521, 0.0002, 1 and 6.1×10^{-5}, respectively, and for the IM-resistant cell data are 0.3451, 0.0988, 1.1×10^{-11} and 0.0001, respectively. (**b**) The frequency of transitions and transversions at codon positions 1 and 2. Transitions: upper (dark grey), lower (black) critical values and expected (dashed) values are given under H_0 for an equal distribution of transitions over codon positions 1 and 2. The number of transitions at codon positions 1 and 2 are shown in light grey. The *P*-value for equal distribution corresponding to codon positions 1 and 2 for the IM-naïve cell data are 0.0810, and for the IM-resistant cell data are 0.0571. Transversions: upper (dark grey) and lower (black) critical values and expected (dashed) values are given under H_0 for an equal distribution of transversions over codon positions 1 and 2. The observed number of transitions at codon positions 1 and 2 are shown in dark grey. The *P*-value for equal distribution corresponding to codon positions 1 and 2 for the IM-naïve cell data are 0.8318, and for the IM-resistant cell data are 0.5052. (**c**) Comparison of observed frequencies of each mutation type at codon positions 1 and 2 with unselected regions of the human genome. The grey bar represents the expected number of each mutation type in the proportions observed in unselected regions of the human genome. The observed number of mutations in the first position are given by the black square and the observed number in the second position by the white triangle.

pattern and profound selection. The lack of position 3 mutations is remarkable, as mutations at the third 'wobble' position are least likely to alter amino acid sequences. This implies that the pathogenicity of the mutator may be linked to its ability to introduce changes at codon positions that are most likely to alter the TK amino acid sequence.

The mutational signature observed in IM-resistant CP-CML patient cells and in cultured CD34$^+$ CML cells was similar. This finding suggests that mutant stem/progenitors arise in an IM therapy-independent manner, but give rise to progeny that acquire a growth advantage due to drug selection. The over-representation of A-to-T transversions in IM-naïve and -resistant cells is a known feature of gliomas, but not of other solid tumours or other haematological malignancies. Note that the mutations characterized were derived from multiple positions across the *BCR–ABL1* TK domain sequence. This makes it unlikely that the observed mutational signature results from the operation of a selection pressure on an as yet undefined primitive CML cell compartment.

It is interesting that the CP-CML mutational signature is strikingly different to that observed in Ph$^+$ B-lymphoid blast crisis, which is characterized by a predominance of G-to-A and C-to-T transitions that are attributed to the activity of activation-induced cytidine deaminase.[8] Activation-induced cytidine deaminase expression was not detected in the CML myeloid cell line (K562) by western blot analysis (Supplementary Figure 1), providing further evidence that the CP-CML mutator is not activation-induced cytidine deaminase and indicating that the transition from CP (myeloid) to blast crisis (B-lymphoid) is accompanied by a switch in the operative mutator.

In addition to providing an insight into the mechanism of somatic point mutation generation in CP-CML, the distinct mutational landscape of the *BCR–ABL1* TK domain in primitive CML cells suggests an approach for anticipating emergent TKI resistance. Patients with a predominance of progenitor clones displaying the 'mutator phenotype' are predicted to experience an overall higher frequency of functional resistance to TKI therapy and may consequently benefit from individualized treatment regimens.

In summary, we describe in this study the use of a mathematical model to compare *BCR–ABL1* TK domain mutations of IM-naïve CD34$^+$ cells and IM-resistant cells from a large number of CP-CML patients. This analysis revealed a distinct and non-random pattern of mutations with hot spots in codons 1 and 2 suggesting the activity of a unique mutator in primitive CP-CML cells.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

The authors thank members of the Leukemia/Bone Marrow Transplant Program of British Columbia for facilitating access to CML patients' samples and clinical data; the Terry Fox Laboratory FACS Facility for assistance in cell sorting; the Stem Cell Assay Laboratory for cell processing and cryopreservation of CML samples; K Saw and K Lambie for excellent technical assistance; and Prof Richard I Christopherson (University of Sydney) and Prof Michael Neuberger (MRC Laboratory of Molecular Biology, Cambridge) for useful discussion. This work was supported in part by grants from the Canadian Cancer Society Research Institute (to XJ, AE and CE), the Cancer Research Society, the Leukemia & Lymphoma Society of Canada and the Canadian Cancer Society (to XJ) with core infrastructure support from the British Columbia Cancer Foundation. X Jiang is Michael Smith Foundation for Health Research Scholar.

H Grant[1,13], X Jiang[2,13], J Stebbing[3], L Foroni[4], C Craddock[5], M Griffiths[6], RE Clark[7], S O'Brien[8], JS Khorashad[4], G Gerrard[4], L Wang[7], JAE Irving[8], M Wang[9], L Karran[10], MJS Dyer[10], D Forrest[11], K Page[1], CJ Eaves[2] and A Woolfson[12]
[1]*Department of Mathematics, University College London, London, UK;*
[2]*Terry Fox Laboratory, British Columbia Cancer Agency and Department of Medical Genetics, University of British Columbia, Vancouver, BC, Canada;*
[3]*Department of Medical Oncology, Centre for Haematology, Imperial College, London, UK;*
[4]*Department of Medicine, Centre for Haematology, Imperial College, London, UK;*
[5]*Centre for Clinical Haematology, Queen Elizabeth Hospital, Birmingham, UK;*
[6]*West Midlands Regional Genetics Laboratory, Birmingham Women's NHS Foundation Trust, Birmingham UK;*
[7]*Royal Liverpool University Hospital, University of Liverpool, Liverpool, UK;*
[8]*Northern Institute for Cancer Research, Newcastle University Medical School, Newcastle upon Tyne, UK;*
[9]*Division of Protein and Nucleic Acid Chemistry, MRC Laboratory of Molecular Biology, Cambridge, UK;*
[10]*MRC Toxicology Unit, University of Leicester, Hodgkin Building, Leicester, UK;*
[11]*Leukaemia/BMT Program of British Columbia, Division of Hematology, British Columbia Cancer Agency, Vancouver, BC, Canada and*
[12]*Discovery Medicine Oncology, Bristol-Myers Squibb, Princeton, NJ, USA*
E-mail: ceaves@bccrc.ca
[13]*These authors contributed equally to this work.*

References

1 Goldman JM, Melo JV. Chronic myeloid leukemia—advances in biology and new approaches to treatment. *N Engl J Med* 2003; **349**: 1451–1464.
2 Druker BJ, Tamura S, Buchdunger E, Ohno S, Segal GM, Fanning S *et al.* Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr-Abl positive cells. *Nat Med* 1996; **2**: 561–566.
3 Apperley JF. Part I: mechanisms of resistance to imatinib in chronic myeloid leukaemia. *Lancet Oncol* 2007; **8**: 1018–1029.
4 Jiang X, Saw KM, Eaves A, Eaves C. Instability of BCR-ABL gene in primary and cultured chronic myeloid leukemia stem cells. *J Natl Cancer Inst* 2007; **99**: 680–693.
5 Jiang C, Zhao Z. Mutational spectrum in the recent human genome inferred by single nucleotide polymorphisms. *Genomics* 2006; **88**: 527–534.
6 Bofkin L, Goldman N. Variation in evolutionary processes at different codon positions. *Mol Biol Evol* 2007; **24**: 513–521.
7 Shah NP, Tran C, Lee FY, Chen P, Norris D, Sawyers CL. Over riding imatinib resistance with a novel ABL kinase inhibitor. *Science* 2004; **305**: 399–401.
8 Klemm L, Duy C, Iacobucci I, Kuchen S, von Levetzow G, Feldhahn N *et al.* The B cell mutator AID promotes B lymphoid blast crisis and drug resistance in chronic myeloid leukemia. *Cancer Cell* 2009; **16**: 232–245.
9 Kulinskaya E. On two-sided *P*-values for nonsymmetric distributions. *Arxiv* 2008; **0810**: 2124.

Supplementary Information accompanies the paper on the Leukemia website (http://www.nature.com/leu)