Dual targeting of p53 and c-MYC selectively eliminates leukaemic stem cells

Sheela A. Abraham[1]*, Lisa E. M. Hopcroft[1]*, Emma Carrick[2,3], Mark E. Drotar[1], Karen Dunn[1], Andrew J. K. Williamson[2,3], Koorosh Korfi[1,4], Pablo Baquero[4], Laura E. Park[1], Mary T. Scott[1], Francesca Pellicano[1], Andrew Pierce[2,3], Mhairi Copland[1], Craig Nourse[4], Sean M. Grimmond[5], David Vetrie[4], Anthony D. Whetton[2,3,6]* & Tessa L. Holyoake[1]*

[1]Paul O'Gorman Leukaemia Research Centre, Institute of Cancer Sciences, University of Glasgow, Gartnavel General Hospital, 1053 Great Western Road, Glasgow G12 0YN, UK.

[2]Stem Cell and Leukaemia Proteomics laboratory, University of Manchester, Manchester M20 3LJ, UK.

[3]Manchester Precision Medicine Institute, University of Manchester, Manchester M20 3LJ, UK.

[4]Institute of Cancer Sciences, University of Glasgow, Garscube Estate, Glasgow G61 1QH, UK.

[5]University of Melbourne Centre for Cancer Research, University of Melbourne, Victoria 3010 Australia.

[6]Stoller Biomarker Discovery Centre, University of Manchester, Manchester M20 3LJ, UK.

*These authors contributed equally to this work.

Chronic myeloid leukaemia (CML) arises after transformation of a haemopoietic stem cell (HSC) by the protein-tyrosine kinase BCR–ABL1. Direct inhibition of BCR–ABL1 kinase has revolutionized disease management, but fails to eradicate leukaemic stem cells (LSCs), which maintain CML. LSCs are independent of BCR–ABL1 for survival, providing a rationale for identifying and targeting kinase-independent pathways. Here we show—using proteomics, transcriptomics and network analyses—that in human LSCs, aberrantly expressed proteins, in both imatinib-responder and non-responder patients, are modulated in concert with p53 (also known as TP53) and c-MYC regulation. Perturbation of both p53 and c-MYC, and not BCR–ABL1 itself, leads to synergistic cell kill differentiation, and near elimination of transplantable human LSCs in mice, while sparing normal HSCs. This unbiased systems approach targeting connected nodes exemplifies a novel precision medicine strategy providing evidence that LSCs can be eradicated.

BCR-ABL1 is a chimaeric oncogene arising from the t(9;22)(q34;q11) chromosomal translocation. The resultant protein tyrosine kinase (PTK) drives signalling events[1] and transforms HSCs. BCR–ABL1 activity in HSCs causes CML, which if untreated, is fatal.

TK inhibitors (TKIs), such as imatinib mesylate, are standard CML treatment and have improved survival, justifying the use of single-target therapies[2]. However, these drugs do not kill the LSCs that maintain the disease[3], resulting in ever-increasing costs to sustain remissions. TKI discontinuation in the best 10–20% of TKI responders led to relapse rates of 50–60%, reinforcing the need to understand and target CML LSCs[4] with curative therapies. Recent studies suggest that LSC survival is BCR–ABL1-kinase independent[5] and BCR–ABL1 has functionality beyond PTK activity, explaining the shortcomings of TKIs[6].

We have applied systems biology approaches to patient material to identify key protein networks that perpetuate the CML phenotype, aiming to elucidate potentially curative therapeutic options. Using unbiased transcriptomic and proteomic analyses, the transcription factors p53 and c-MYC are identified as having defining roles in CML LSC survival. We demonstrate an integral relationship between p53 and c-MYC in the maintenance of CML, and importantly, a potential therapeutic advantage they have as drug targets over BCR–ABL1 for eradication of CML LSCs.

p53 and c-MYC mediate the CML network

To interrogate perturbations in BCR–ABL1 signalling of potential therapeutic value, isobaric-tag mass spectrometry (MS) was used to compare treatment-naive CML and normal $CD34^+$ cells. Fifty-eight proteins were consistently deregulated in three CML samples (Methods and Supplementary Table 1). Dijkstra's algorithm[7] and the MetaCore knowledge base (https://portal.genego.com/) were used to identify p53 and c-MYC as central hubs (Supplementary Table 2) in a CML network of 30 proteins (Fig. 1a) predominantly downstream of the transcription factors, with significant enrichment for p53/c-MYC targets (Fisher's exact test, $P = 0.001$). While the majority of proteins downstream of p53 were downregulated, those downstream of c-MYC included proteins up- or downregulated in CML, in keeping with c-MYC acting as an activator and repressor of gene transcription[8]. The deregulated network suggests an altered dependency on p53 and c-MYC in CML $CD34^+$ cells.

This data set represents the first—to our knowledge—relative quantitative comparison of CML to normal $CD34^+$ cells using MS. Importantly, CML-initiating cells reside within the $CD34^+CD38^-Lin^-$ subpopulation and may differ to bulk $CD34^+$ cells. To substantiate the CML proteome observations and investigate regulation in LSCs, we examined relevant, primary CML transcriptomic data. Network protein levels correlated well with respective gene levels, in both

LSCs (four independent data sets; Fig. 1b and Extended Data Fig. 1a–c) and $CD34^+$ progenitors (Extended Data Fig. 1d, e). Correlations were stronger for the 30 network candidates compared to all 58 deregulated proteins; seven data sets showed significant gain in r^2 for network candidates (Extended Data Fig. 1a, d). The mutual information (MI) of proteomic/transcriptomic data for network proteins was significantly greater than random (Fig. 1c and Extended Data Fig. 1b, e). This consistent messenger RNA/protein correspondence, in both progenitors and LSCs, confirmed that the network was transcriptionally regulated, compatible with c-MYC and p53 function.

p53 and c-MYC have key roles in oncogenesis and appear in many cancer networks. To distinguish true regulatory effectors, we assessed the bias towards outgoing versus incoming signalling ($\text{degree}_{\text{out}}/\text{degree}_{\text{in}}$ ($d_{\text{out}}/d_{\text{in}}$)) or p53 and c-MYC. We generated networks from deregulated proteins in (1) primary MS data sets[9–11], (2) cell lines transduced with oncogenic PTKs driving haematological malignancies[12], and from (3) 50 randomly generated protein sets. Our network falls outside the expected random distribution and no other data set exhibits greater downstream bias for p53 and c-MYC (Fig. 1d). These data support a novel network in, and unique to, CML, centred on p53 and c-MYC.

Validation of network candidates

The CML network revealed well-characterized p53/c-MYC targets and proteins not previously associated with CML pathogenesis (Supplementary Table 3). To validate proteomic predictions (Fig. 1a), gelsolin, CIP2A (also known as KIAA1524), UCHL1, aldose reductase, p53 and c-MYC were assessed using western blotting and immunofluorescence (Fig. 2a, b). Protein expression of gelsolin, CIP2A, UCHL1 and aldose reductase were consistent with CML network predictions (Fig. 2a). Immunofluorescence was confirmatory, highlighting the dramatic difference in CIP2A expression between normal and CML cells, and the intracellular localization of gelsolin and aldose reductase (Fig. 2b). CML cells also expressed increased c-MYC and decreased p53 levels (Fig. 2a, b), correlating well with appropriate modulation of downstream targets. We therefore hypothesized that simultaneous p53 activation and c-MYC inhibition would kill LSCs. To assess dual hub requirement for CML survival, lentiviral short hairpin RNA (shRNA) constructs (Extended Data Fig. 2a) were employed. Knockdown of HDM2 (also known as MDM2; E3 ligase/negative regulator for p53), c-MYC, or both, in CML $CD34^+$ cells reduced

viability and enhanced apoptosis; the combined effects were synergistic. In colony-forming cell (CFC) assays, effects were more dramatic with single or combined knockdown, strengthening the hypothesis that p53 and c-MYC are critical for the survival of CML cells (Fig. 2c–e and Extended Data Fig. 2b, c). We then investigated synergistic interactions between p53 and c-MYC, testing clinically tractable inhibitors.

RITA and CPI-203 synergize to drive CML CD34$^+$ cell kill

To target identified hubs, we selected RITA (also known as NSC 652287),which binds p53 and blocks its degradation, and CPI-203, a bromodomain and extra terminal protein (BET) inhibitor hindering transcription by disrupting chromatin-dependent signal transduction[13,14]. As anticipated[15], c-MYC was downregulated 8 h after CPI-203 treatment (Extended Data Fig. 3a). CPI-203 also reduced p53 at 8 h. RITA subtly increased p53 by 8 h, further enhanced by 24–48 h. Dasatinib (Das) at 150 nM (a concentration achievable in patients to fully inhibit BCR–ABL[16]) gradually reduced p53 levels and inhibited phosphorylation of STAT5 as previously observed[17,18]. RITA with CPI-203 for 8 h reduced both p53 and c-MYC, suggesting a dominant effect of CPI-203 at this early time point, but by 48 h markedly increased p53 (Extended Data Fig. 3a–c).

RITA or CPI-203 treatment of CML CD34$^+$ cells for 72 h reduced viability in a concentration-dependent manner and induced significant apoptosis; combining drugs resulted in further significant reductions in these parameters (Fig. 3a, b and Extended Data Fig. 3d). Labelling with the cell division tracker carboxyfluorescein succinimidyl ester (CFSE) and CD34 antibody showed that as CML cells divided in the presence of CPI-203, there was clear and rapid loss of CD34 expression not seen with RITA (Fig. 3c), suggesting that c-MYC (a predominant target of CPI-203; ref. 15) inhibition induces differentiation of CML CD34$^+$ cells. Differentiation was further suggested by skewing of the morphology, size and number of CFCs (Fig. 3d). RITA decreased CFCs but did not affect colony types (Fig. 3d, e). By inducing apoptosis and differentiation, the drugs may synergize and enhance elimination of CML. By measuring drug–dose response[19], combination therapy was potently synergistic with combination indicies (CI) ranging from 0.07 to 0.34 (Fig. 3a). Nutlin-3a (Nut), another HDM2 inhibitor, produced similar results. The effects of RITA and Nut were p53 dependent, as K562 cells lacking p53 were non-responsive (Extended Data Figs 3e and 4a, b). Since CPI-203 ± RITA reduced p53 at the early

time point, sequential inhibition of HDM2 and c-MYC was tested using chemical and genetic approaches. Neither inhibition of HDM2 before c-MYC nor vice versa improved cell kill compared with simultaneous knockdown or drug inhibition, or compared with nilotinib (Nil) (Extended Data Fig. 4c, d).

CML patients receive a TKI, irrespective of response. We therefore assessed RITA and CPI-203 effects in imatinib-mesylate-pre-treated CML $CD34^+$ cells. Imatinib mesylate neither ameliorated nor enhanced the efficacy of RITA and/or CPI-203 (Extended Data Fig. 4e).

RITA and CPI-203 eliminate LSCs

The CML network suggested an altered dependency on p53/c-MYC signalling. We therefore hypothesized that normal cells may be less susceptible to the drug combination. Treatment of normal $CD34^+$ cells with single agents or combinations had no significant effects on cell counts. However, increased apoptosis was observed at higher CPI-203 concentrations (2 or 5 μM) and with the highest combination (RITA 25 nM, CPI-203 5 μM; Extended Data Fig. 5a, b). In CML cells, apoptosis was observed with all four CPI-203 and combination concentrations (Fig. 3b), confirming a therapeutic window.

To confirm the *in silico* results that led to the prediction (Fig. 1b, c and Extended Data Fig. 1c) that an altered dependency on p53 and c-MYC extended to primitive LSCs, we exposed CML LSCs to the drug combination. LSCs were defined as either $CFSE^{max}$ or $CD34^+CD38^-$. As shown previously[16], in comparison to untreated control, over 5 days the $CFSE^{max}$ population persisted in response to Das and Nil, but was significantly reduced by CPI-203 alone and by combination treatment (Figs 3c and 4a, b). Over 72 h, RITA with CPI-203 was also effective in synergistically eliminating residual $CD34^+CD38^-$ cells (CI = 0.3–0.8; Extended Data Fig. 5c).

HSCs and LSCs are most stringently defined by their engraftment capacity at 16 weeks. We exposed CML $CD34^+$ cells to RITA, CPI-203, the combination, or Das for 48 h before transplantation into sublethally irradiated NSG mice (Extended Data Fig. 5d). Human $CD45^+$ cells were detectable in peripheral blood at 8, 12 and 16 weeks and in bone marrow at 16 weeks post-transplantation. Das had no significant effect on NSG-repopulating CML LSCs, representing the most primitive long-term engrafting cells. In contrast RITA, CPI-203, and the combination reduced engraftment as indicated by decreased $CD45^+$, $CD34^+$, $CD33^+$, $CD11b^+$, $CD19^+$ and $CD14^+$ cells (Extended Data Fig. 5e, f). Using a CML sample known to engraft both

BCR–ABL1$^+$ and BCR–ABL1$^-$ cells by double fusion fluorescence *in situ* hybridization (D-FISH), there was a marked decrease in the long-term-engrafting potential of RITA-, CPI-203- or combination-treated leukaemic cells, with no significant effect on non-leukaemic populations (Fig. 4c, d and Extended Data Fig. 5g, h). Experiments using cord blood CD34$^+$ cells confirmed the selectivity of RITA and CPI-203 for BCR–ABL1$^+$ versus BCR–ABL1$^-$ stem cells (Fig. 4d).

Mechanism of LSC kill and clinical scope

To understand the mechanism(s) underlying reduction of CML stem and progenitor cells in response to RITA and CPI-203, RNA sequencing (RNA-seq) was performed. Of the 12,248 genes sequenced, 2,134 were identified as synergistically modulated by combination treatment; 166 demonstrated extreme synergy (Fig. 5a). Moreover, 81% of the genes differentially expressed in response to the combination were deregulated in the same direction with RITA or CPI-203 ($\chi^2(1) = 891.93$, $P < 0.01$). While transcriptional responses to RITA or CPI-203 were enriched for p53/apoptosis or c-MYC/differentiation (not found with Nil), respectively, the combination induced enhanced or additional enrichment of these molecular signatures and pathways (Fig. 5b, Extended Data Fig. 6a–c and Supplementary Tables 5–7). Furthermore, stem/progenitor markers CD34 and CD133 were dramatically downregulated by CPI-203 and the combination, but not by RITA or Nil (Extended Data Fig. 7b). The enrichment in the p53/apoptosis and c-MYC/differentiation pathways paralleled the *in vitro* phenotypic effects observed. Limited overlap in gene membership of the signatures identified *in silico* demonstrates that distinct molecular components contribute to single and combined drug responses (Fig. 5c).

CML stem cell persistence is an issue for all CML patients, however, many also exhibit or acquire TKI resistance or demonstrate a more aggressive clinical phenotype[20]. These represent patients in whom novel agents targeting p53 and c-MYC would first be tested. To investigate whether the deregulated p53/c-MYC network is present in both TKI-responder (TKI-R) and TKI-non-responder (TKI-NR) patients[21], and in more advanced forms of CML[20], we considered data from CD34$^+$ CML cells derived from suitable patient cohorts. Transcriptional expression of the network components was highly correlated across all CML versus normal comparisons, irrespective of TKI response or clinical phenotype (Fig. 5d). In keeping with these *in silico* data, CD34$^+$ cells from a TKI-NR patient showed high levels of apoptosis after treatment with RITA

and/or CPI-203 (Extended Data Fig. 7c), suggesting that these drugs should be of therapeutic value for such patients.

RG7112/7388 and CPI-203/0610 therapy

To progress the drug combination towards the clinic, we used complementary preclinical mouse models and introduced RG7112 and RG7388, both HDM2 inhibitors[22], and CPI-0610, a BET inhibitor[23]; drugs already advanced in clinical trials in humans. In the SCL-tTA-*BCR-ABL* double transgenic (DTG) leukaemia mouse model BCR–ABL1, driven from the stem cell promoter (SCL), is inducible in HSCs by tetracycline withdrawal (tTA), resulting in a transplantable CML-like disease with increased myeloid counts and splenomegaly[16,17,24].

After irradiation, C57BL/6 $CD45.1^+$ mice were used as recipients and $CD45.2^+$ mice as DTG bone marrow donors (Extended Data Fig. 8a). After transplantation (to synchronize leukaemia development and assess transplantable LSCs), CML was induced. In mice and/or rats, RG7112 at 50–200 mg kg^{-1} and CPI-0610 at 15–60 mg kg^{-1} have demonstrated on-target effects in tumours[22,23]. Excellent tolerability was achieved with modest doses of RG7112 (50 mg kg^{-1} once daily) and CPI-0610 (15 mg kg^{-1} twice daily, both for 4 weeks), selected to demonstrate synergy. White blood cell and neutrophil counts returned to non-leukaemic control levels with the drug combination, but not with single treatments (Fig. 6a and Extended Data Fig. 8b). While CPI-0610, Nil, and the combination significantly reduced spleen size (Fig. 6b), only the combination significantly reduced donor leukaemic $CD45.2^+$ cells in the bone marrow (protected by the niche), while simultaneously allowing recovery of host normal $CD45.1^+$ cells; CD45.1:CD45.2 ratio changed from 20:80 (untreated) to 40:60 (combination) (Extended Fig. Data 8c). At the stem cell level, donor leukaemic $Lin^-Sca\text{-}1^+c\text{-}Kit^+$ ($CD45.2^+$ LSK) cells were reduced by >60% by the combination, while host LSKs were unaffected. None of the single arms reduced LSKs (Fig. 6b–d), supporting the synergistic effects demonstrated *in vitro*.

To confirm that these therapeutic *in vivo* effects extended to human CML, two cohorts of sublethally irradiated NSG mice were transplanted with independent CML $CD34^+$ samples and treated with RG7388 and CPI-203 (75–100 mg kg^{-1} and 6–7.5 mg kg^{-1}, respectively) for 3–4 weeks. Of the single agents, only CPI-203 showed a consistent effect. The drug combination, however, eliminated 95% of Ph^+CD45^+ and 88% of $CD45^+CD34^+$ subsets (Fig. 6e, f). These results were significant as compared to vehicle ($P < 0.001$) or Nil ($P = 0.0016$ ($CD45^+$);

$P = 0.0047$ ($CD45^+CD34^+$)) and as compared to RG7388 ($P = 0.0017$ ($CD45^+$); $P = 0.0004$ ($CD45^+CD34^+$)) or CPI-203 ($P = 0.0046$ ($CD45^+$); $P = 0.0008$ ($CD45^+CD34^+$)), respectively (Fig. 6e, f), again suggesting a high degree of synergy.

Discussion

This work demonstrates the potential of unbiased, systems approaches to uncover new therapeutic options by analysing primitive stem cell subsets from primary material. We found that p53 and c-MYC work together with BCR–ABL1 to shape leukaemia stem cell phenotype and show that modulation of both p53 and c-MYC is critical to drive synergistic enhancement of apoptosis and differentiation seen *in vitro*, *in vivo*, at the stem cell level and the molecular level by RNA-seq.

p53 and c-MYC have individually been identified as proteins in CML pathobiology[17,25–30] and cancer[31–33], but have not previously been considered for dual targeting. In recent CML studies, enhanced LSC kill converged on p53 as the mediator of apoptosis[17,25,34]. CML LSCs are also susceptible to enhancement or depletion of c-MYC. After deletion of the E3 ligase FBXW7, c-MYC increases with p53, resulting in cell cycle entry and p53-dependent apoptosis[29,30]. However, FBXW7 may not represent a viable drug target based on its role in haematopoiesis, tissue stem cells and, importantly, as a tumour suppressor. Currently, there is interest in drugging the spliceosome machinery, particularly for MYC-driven cancers, however, a therapeutic window remains to be established and exploited with well-tolerated agents[35,36]. BET inhibition is a rapidly expanding market with multiple agents in phase 1–2 development. These agents are well tolerated and demonstrate efficacy in haematological malignancies. Resistance to BET inhibitors evolves through epigenetic mechanisms[37], however, our combination approach will be less susceptible to resistance.

Over the last decade CML has been transformed from a fatal cancer to a manageable disease with lifelong therapy. Despite recognition that LSCs prevent cure, the paradigm established by TKIs means that novel drugs must be safe, supported by a clear therapeutic window and easy to administer. As a result, few preclinical studies have reached the clinic and trials fail owing to toxicity and poor recruitment[24,38]. CML is often regarded as a simple cancer, driven solely by BCR–ABL1, yet we do not understand why targeting BCR–ABL1 does not eradicate LSCs, nor cure CML. Our work shows that BCR–ABL1 reprograms potent

oncoproteins and tumour suppressors, to establish a signalling network that underlies the propagation of CML. Critically, we found that simultaneous perturbation of p53 and c-MYC, mechanistically driving modulation of p53, apoptosis, c-MYC and differentiation pathways, improved selective kill of LSCs as compared to TKIs. Nil was ineffective against these pathways, potentially explaining why TKIs are not sufficient to cure CML. The fact that the aberrant network was similarly regulated in TKI-Rs, TKI-NRs, and patients with aggressive and indolent CML, coupled with the availability of well-tolerated oral agents in an advanced stage of development, now offers an entirely novel approach for the treatment of CML, with the therapeutic potential to address CML LSC persistence and improve outcome for CML patients.

Received 8 March; accepted 26 April 2016.

Published online XX 2016.

1. O'Hare, T., Zabriskie, M. S., Eiring, A. M. & Deininger, M. W. Pushing the limits of targeted therapy in chronic myeloid leukaemia. *Nature Rev. Cancer* **12,** 513–526 (2012).

2. Druker, B. J. *et al.* Activity of a specific inhibitor of the BCR-ABL tyrosine kinase in the blast crisis of chronic myeloid leukemia and acute lymphoblastic leukemia with the Philadelphia chromosome. *N. Engl. J. Med.* **344,** 1038–1042 (2001).

3. Graham, S. M. *et al.* Primitive, quiescent, Philadelphia-positive stem cells from patients with chronic myeloid leukemia are insensitive to STI571 *in vitro.* *Blood* **99,** 319–325 (2002).

4. Mahon, F. X. *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.* **11,** 1029–1035 (2010).

5. Corbin, A. S. *et al.* Human chronic myeloid leukemia stem cells are insensitive to imatinib despite inhibition of BCR-ABL activity. *J. Clin. Invest.* **121,** 396–409 (2011).

6. Neviani, P. *et al.* PP2A-activating drugs selectively eradicate TKI-resistant chronic myeloid leukemic stem cells. *J. Clin. Invest.* **123,** 4144–4157 (2013).

7. Dijkstra, E. W. A note on two problems in connexion with graphs. *Numer. Math.* **1,** 269–271 (1959).

8. Van Dang, C. & McMahon, S. B. Emerging concepts in the analysis of transcriptional targets of the MYC oncoprotein: are the targets targetable? *Genes Cancer* **1,** 560–567 (2010).

9. Boichenko, A. P. *et al.* A panel of regulated proteins in serum from patients with cervical intraepithelial neoplasia and cervical cancer. *J. Proteome Res.* **13,** 4995–5007 (2014).

10. Liu, Y. *et al.* Glycoproteomic analysis of prostate cancer tissues by SWATH mass spectrometry discovers N-acylethanolamine acid amidase and protein tyrosine kinase 7 as signatures for tumor aggressiveness. *Mol. Cell. Proteomics* **13,** 1753–1768 (2014).

11. Sutton, C. W. *et al.* Quantitative proteomic profiling of matched normal and tumor breast tissues. *J. Proteome Res.* **9,** 3891–3902 (2010).

12. Pierce, A. *et al.* Eight-channel iTRAQ enables comparison of the activity of six leukemogenic tyrosine kinases. *Mol. Cell. Proteomics* **7,** 853–863 (2008).

13. Issaeva, N. *et al.* Small molecule RITA binds to p53, blocks p53–HDM-2 interaction and activates p53 function in tumors. *Nature Med.* **10,** 1321–1328 (2004).

14. Wong, C. *et al.* The bromodomain and extra-terminal inhibitor CPI203 enhances the antiproliferative effects of rapamycin on human neuroendocrine tumors. *Cell Death Dis.* **5,** e1450 (2014).

15. Mertz, J. A. *et al.* Targeting MYC dependence in cancer by inhibiting BET bromodomains. *Proc. Natl Acad. Sci. USA* **108,** 16669–16674 (2011).

16. Hamilton, A. *et al.* Chronic myeloid leukemia stem cells are not dependent on Bcr-Abl kinase activity for their survival. *Blood* **119,** 1501–1510 (2012).

17. Li, L. *et al.* Activation of p53 by SIRT1 inhibition enhances elimination of CML leukemia stem cells in combination with imatinib. *Cancer Cell* **21,** 266–281 (2012).

18. Shah, N. P. *et al.* Transient potent BCR-ABL inhibition is sufficient to commit chronic myeloid leukemia cells irreversibly to apoptosis. *Cancer Cell* **14,** 485–493 (2008).

19. Chou, T. C. Drug combination studies and their synergy quantification using the Chou-Talalay method. *Cancer Res.* **70,** 440–446 (2010).

20. Yong, A. S., Szydlo, R. M., Goldman, J. M., Apperley, J. F. & Melo, J. V. 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* **107,** 205–212 (2006).

21. McWeeney, S. K. *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* **115,** 315–325 (2010).

22. Tovar, C. *et al.* MDM2 small-molecule antagonist RG7112 activates p53 signaling and regresses human tumors in preclinical cancer models. *Cancer Res.* **73,** 2587–2597 (2013).

23. Albrecht, B. K. *et al.* Identification of a benzoisoxazoloazepine inhibitor (CPI-0610) of the bromodomain and extra-terminal (BET) family as a candidate for human clinical trials. *J. Med. Chem.* **59,** 1330–1339 (2016).

24. Zhang, B. *et al.* Effective targeting of quiescent chronic myelogenous leukemia stem cells by histone deacetylase inhibitors in combination with imatinib mesylate. *Cancer Cell* **17,** 427–442 (2010).

25. Hurtz, C. *et al.* BCL6-mediated repression of p53 is critical for leukemia stem cell survival in chronic myeloid leukemia. *J. Exp. Med.* **208,** 2163–2174 (2011).

26. Lanza, F. & Bi, S. Role of p53 in leukemogenesis of chronic myeloid leukemia. *Stem Cells* **13,** 445–452 (1995).

27. Gómez-Casares, M. T. *et al.* MYC antagonizes the differentiation induced by imatinib in chronic myeloid leukemia cells through downregulation of p27(KIP1). *Oncogene* 32, 2239–2246 (2013).

28. Winter, G. E. *et al.* Systems-pharmacology dissection of a drug synergy in imatinib-resistant CML. *Nature Chem. Biol.* **8,** 905–912 (2012).

29. Takeishi, S. *et al.* Ablation of Fbxw7 eliminates leukemia-initiating cells by preventing quiescence. *Cancer Cell* **23,** 347–361 (2013).

30. Reavie, L. *et al.* Regulation of c-Myc ubiquitination controls chronic myelogenous leukemia initiation and progression. *Cancer Cell* **23,** 362–375 (2013).

31. Dang, C. V. MYC, metabolism, cell growth, and tumorigenesis. *Cold Spring Harb. Perspect. Med.* **3,** a014217 (2013).

32. Vousden, K. H. & Prives, C. Blinded by the light: the growing complexity of p53. *Cell* **137,** 413–431 (2009).

33. Wade, M., Li, Y. C. & Wahl, G. M. MDM2, MDMX and p53 in oncogenesis and cancer therapy. *Nature Rev. Cancer* **13,** 83–96 (2013). </jrn>

34. Zhang, H., Li, H., Ho, N., Li, D. & Li, S. Scd1 plays a tumor-suppressive role in survival of leukemia stem cells and the development of chronic myeloid leukemia. *Mol. Cell. Biol.* **32,** 1776–1787 (2012).

35. Hsu, T. Y. *et al.* The spliceosome is a therapeutic vulnerability in MYC-driven cancer. *Nature* **525,** 384–388 (2015).

36. Koh, C. M. *et al.* MYC regulates the core pre-mRNA splicing machinery as an essential step in lymphomagenesis. *Nature* **523,** 96–100 (2015).

37. Fong, C. Y. *et al.* BET inhibitor resistance emerges from leukaemia stem cells. *Nature* **525,** 538–542 (2015). </jrn>

38. Prost, S. *et al.* Erosion of the chronic myeloid leukaemia stem cell pool by PPARγ agonists. *Nature* **525,** 380–383 (2015).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank all CML patients and UK haematology departments who contributed samples; A. Hair for sample processing; J. Cassels for cell sorting; A. Michie and V. Helgason for assisting with the *in vivo* research and providing cord blood samples; C. Wells, P. Bailey and J. Cole for discussions regarding the RNA-seq analysis. We acknowledge the Cancer Research UK (CR-UK) Glasgow Centre (C596/A18076) and the CR-UK Beatson Institute (C596/A17196) for providing animal care and housing facilities. We acknowledge Constellation Pharmaceuticals for providing CPI-203, CPI-0610 and part funding M.E.D., Roche for providing RG7112, RG7388 and part funding M.E.D. and the SPIRIT Trials Management Group for access to CML samples. This study was supported by the Glasgow and Manchester Experimental Cancer Medicine Centres (ECMC), which are funded by CR-UK and the Chief Scientist's Office (Scotland). We acknowledge the funders who have contributed to this work: MRC stratified medicine infrastructure award (A.D.W.), CR-UK C11074/A11008 (F.P., L.E.M.H., T.L.H.,

A.D.W.); LLR08071 (S.A.A., E.C.); LLR11017 (M.C.); SCD/04 (M.C.); LLR13035 (S.A.A., K.D., A.D.W. and A.P.); LLR14005 (M.T.S., D.V.); KKL690 (L.E.P.); KKL698 (P.B.); LLR08004 (A.D.W., A.P. and A.J.W.); MRC CiC (M.E.D.); The Howat Foundation (fluorescence-activated cell sorting (FACS) support); Friends of Paul O'Gorman (K.D. and FACS support); ELF 67954 (S.A.A.); British Society for Haematology start-up fund (S.A.A.); MR/K014854/1 (K.D.).

Author Contributions A.D.W. and T.L.H. supervised the entire study and research. S.A.A., L.E.M.H., A.D.W. and T.L.H. designed the research, analysed and interpreted data, and wrote the manuscript. S.A.A. conceived the hypothesis, supervised *in vivo* research, prepared samples for proteomic and RNA-seq, performed all *in vitro* work including western blotting, immunofluorescence, cloning and knockdown studies, clonogenic studies, flow cytometry and all mouse *in vivo* studies including tissue processing-FISH preparation and slide interpretation, engraftment determination and analysis of primitive stem cell subsets. L.E.M.H. designed and performed all *in silico* work including global omics handling, integration and analysis (MS, RNA-seq and microarray data); network analyses; correlation/MI calculations; functional enrichment analyses and permutation experiments for calculation of *P* values. E.C. performed proteomic work. A.J.K.W. performed proteomic work and generated relative proteomic quantification. M.E.D. performed virus preparation, prepared drugs for *in vivo* work and assisted with *in vivo* studies. K.D. provided maintenance and care for all mouse colonies and assisted with *in vivo* work. P.B. and L.E.P. provided assistance with *in vivo* studies and D-FISH preparation. F.P. and M.T.S. provided assistance with *in vivo* studies. D.V. provided analysed datasets and analysed/interpreted RNA-seq data. S.M.G. supervised and interpreted RNA-seq. C.N. performed RNA-seq experiments. K.K. and M.C. provided analysed datasets. A.P. supervised proteomic studies. All authors reviewed/edited the manuscript.

Author Information The CML and normal $CD34^+$ mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifiers PXD001502, PXD001503, PXD001504; SCOPE3/SCOPE4 data are also available using PXD001505 and PXD002782 respectively. Transcriptomic data are publicly available via the accession codes E-MTAB-2581, E-MTAB-2508, E-MIMR-17 at ArrayExpress (https://www.ebi.ac.uk/arrayexpress/) and GSE47927, GSE5550, GSE24739, GSE14671 at GEO (http://www.ncbi.nlm.nih.gov/geo/). RNA-seq data (fastq) have been deposited in the European Nucleotide Archive under accession number PRJEB9942. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.L.H. (Tessa.Holyoake@glasgow.ac.uk).

Figure 1 p53 and c-MYC network in CML regulation. a, Network analysis reveals that c-MYC and p53 are central in a putative CML network (n=3 patient samples, n=2 normal samples). **b**, Correlation between proteomic/transcriptomic deregulation in primitive $CD34^+Hst^{lo}Py^{lo}$ (G0) (top two panels); $CD34^+CD38^-$ (second panel from the bottom) and $Lin^-CD34^+CD38^-CD90^+$ CML cells (bottom panel). Filled black circles indicate all

protein/genes; filled red circles indicate network proteins/genes. r_c, full candidate list correlation; r_n, network correlation. **c**, Gene/protein MI for the CML network (red, false discovery rate (FDR) < 0.05; grey, FDR < 0.10); FDR calculated using 10,000 re-samplings (blue histogram). **d**, The d_{out}/d_{in} ratio for p53 and c-MYC in haematological PTK-regulated cell lines (n=3 BaF/3 technical replicates per transfected oncogene), other primary cancers and random protein networks.

Figure 2 Validation of proteomic network. a, Network proteins, p53 and c-MYC western blots using CML and normal $CD34^+$ cells. For gel source data, see Supplementary Fig. 1. **b**, Network proteins validated by immunofluorescence in CML and normal $CD34^+$ cells labelled green or red (far left), with the nucleus stained blue using 4′,6-diamidino-2-phenylindole (DAPI; second left), overlays of images (second right), and three-dimensional fluorescent signal (far right). **c**, CML $CD34^+$ cells after HDM2, c-MYC, or scramble knockdown (n = 6 patient samples). **d**, Apoptosis after knockdown (n = 4 patient samples). **e**, CFCs from knockdown (n = 3 patient samples). Values normalized to scrambled control, mean ± standard error of the mean (s.e.m.) (P values: two-tailed Student's t-test; *P < 0.05, **P < 0.01, ***P < 0.001).

Figure 3 Modulation of p53 and c-MYC demonstrates CML sensitivity. a, **b**, Drug titrations: cell viability with CIs (n = 3 patient samples) (**a**) and apoptosis (n = 3 patient samples) (**b**). **c**, CFSE/CD34-labelled cells. Cell divisions are multicoloured. Representative of n = 3 patient samples. **d**, CFCs from treated cells. Representative of n = 3 patient samples. **e**, Averaged CFCs (n = 3 patient samples). One experiment is represented in **d**, mean ± s.e.m. (P values: two-tailed Student's t-test; *P < 0.05, **P < 0.01, ***P < 0.001).

Figure 4 p53 and c-MYC abrogation in normal and primitive CML cells. a, CFSE/CD34-labelled CML cells. Combo, combination. **b**, Recovery of $CFSE^{max}$ CML cells after 5 days treatment (n = 3 patient samples). **c**, Bone marrow (BM) analyses of human CML, replicated twice (2 patient samples), with a minimum of n = 6 mice per arm. **d**, Bone marrow analyses of human cord blood replicated once, n = 5 mice per arm; mean ± s.e.m. (P values: two-tailed Student's t-test; *P < 0.05, **P < 0.01, ***P < 0.001).

Figure 5 Mechanism and clinical relevance of treatment. a, Molecular synergy for 100 nM RITA, 1 μM CPI-203 and RITA plus CPI-203 (Combo) 24 h treatment (NDC=no drug control; n=3 patient samples per arm); mean (μ) expression of 'all' and 'extreme' synergistic genes

summarized as indicated. FC, fold change. **b**, Enrichment of p53 (far left), apoptosis (second left), c-MYC (second right) and differentiation (far right) Molecular Signatures Database (MSigDB) signatures. Asterisks indicate significant enrichment specific to combination treatment. **c**, Gene membership of three functional signatures. **d**, Comparison of transcriptional profiles of TKI-R, TKI-NR and baseline CML versus normal (left), and aggressive, indolent and baseline CML versus normal (right) for our candidate network (Fig. 1a). FDRs calculated by 10,000 permutations.

Figure 6 Targeting p53 and c-MYC in CML elicits synergistic kill in BCR–ABL1$^+$ LSCs. a, White blood cell (WBC) counts and spleen weights normalized to control (dotted line) (experiments replicated twice, minimum $n = 7$ mice per arm; vehicle, no drug control). Combo, combination treatment. **b**–**d**, Bone marrow stained for CD45.1/2 and further gated on Lin$^-$Sca-1$^+$c-Kit$^+$ (LSK). Drug treatments (experiments replicated twice, minimum $n = 5$ mice per arm). **e**, **f**, NSG mice *in vivo* treatment: bone marrow stained for human Ph$^+$CD45$^+$ (left) and further gated on CD34$^+$ cells (right). **g**, Representative CD34$^+$ dotplots (experiments replicated twice (2 patient samples), minimum $n = 9$ mice per arm); mean ± s.e.m. (*P* values: two-tailed Student's *t*-test; $*P < 0.05$, $**P < 0.01$, $***P < 0.001$).

METHODS

Patient samples

Patient samples (PS) were leukapheresis products taken at time of diagnosis with chronic-phase CML, with written informed consent in accordance with the Declaration of Helsinki and approval of the Greater Glasgow and Clyde National Health Service Trust Institutional Review Board. CD34$^+$ cells were enriched using CliniMACS (Miltenyi Biotec), with stem cell subsets purified by FACS. CML CD34$^+$ samples were cultured in serum-free medium (SFM) supplemented with growth factors as described previously[16]. Normal CD34$^+$ cells were CD34-enriched leukapheresis products or cord blood maintained as described for CML CD34$^+$ samples. All PS and relevant clinical data are summarized in Supplementary Table 4. All *in vitro* work was performed with a minimum of 3 PS (3 biological replicates) unless otherwise indicated. Unless otherwise indicated each PS was analysed as an individual sample replicated once in an experiment.

Cell lines

The HeLa cell line was obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ; originally deposited by ATCC). HeLa cells were subcultured in RPMI 1640 (10% FCS plus 2 mM L-glutamine, 100 units ml^{-1} penicillin and 100 µg ml^{-1} streptomycin (Gibco- Life Technologies)) (passage 3). The K562 cell line (DSMZ) was subcultured in IMDM (10% FCS plus 2 mM L-glutamine, 100 units ml^{-1} penicillin and 100 µg ml^{-1} streptomycin (passage <6)); the cell lines were not authenticated between passage 2–6. Cell lines were mycoplasma negative in DAPI, microbiological culture, RNA hybridization and PCR assays.

CML cytoplasmic preparations for MS

CML and normal $CD34^+$ cells were thawed and cultured overnight as described[24]. Cytoplasmic preparations were prepared using the Active Motif Nuclear Extraction Kit (Active Motif).

Materials

RITA (CAS 213261-59-7; catalogue no. 10006426), Nutlin-3a (CAS 675576-98-4; catalogue no. 18585) (Cayman Chemical) and Nil (CAS 641571-10-0) (Selleck Chemicals) were stored as per manufacturer's instructions. CPI-203 and CPI-0610 were obtained from Constellation Pharmaceuticals and kept as a solid powder at room temperature. Das (Selleck Chemicals) was kept as a stock solution (10 mg ml^{-1}) in dimethylsulfoxide (DMSO; Sigma-Aldrich) and prepared and stored in aliquots at −20 °C. Imatinib mesylate (LC Laboratories) was stored at 100 mM in distilled water at 4 °C. RG7112 and RG7388 were supplied by Roche. A list of all antibodies used is provided in Supplementary Table 8.

Proteomics

Methods used have been described[12]. Twenty micrograms of protein was isobarically tagged (iTRAQ reagent, ABSciex). Peptides were identified by Reverse phase liquid chromatography tandem mass spectrometry (RP-LC-MS/MS) on three different instruments: ABSciex Q-STAR Elite, Thermo LTQ Orbitrap Velos, ABSciex TripleTOF 5600. For the 5600 and Elite, dried peptide fractions were resuspended in 15 µl 3% (v/v) acetonitrile, 0.1% (v/v) formic acid and 20 mM citric acid. For each analysis, a 5 µl peptide sample was loaded onto a nanoACQUITY UPLC Symmetry C18 Trap (5 µm, 180 µm × 20 mm) and separation of the peptides was performed using nanoACQUITY UPLC BEH C18 Column (1.7 µm, 75 µm × 250 mm). For the Orbitrap, 10% of the peptide sample was loaded onto Acclaim PepMap µ-Precolumns, analytical

separation of the peptides was performed using Acclaim PepMap RSLC C18 Columns. Data were acquired using the information-dependent acquisition (IDA) protocol.

Elite and 5600 data were processed by a 'thorough' search against the UniProtKB/SwissProt human database containing 532,146 sequence entries using ProteinPilot Software 4.1, revision number 460, Paragon Algorithm 4.0.0.0. Orbitrap data were analysed using Proteome Discoverer 1.3. The data were searched using the MASCOT node of Proteome Discoverer with the UniProtKB database (release 2011_11). The proteins observed in the three data sets demonstrated that using multiple instruments enhanced coverage (Extended Data Fig. 1).

Analysis and integration of MS proteomic data

MS data sets were filtered for peptides observed in all channels (one normal sample was removed from all experiments due to poor labelling). Deregulated proteins were identified using a threshold of mean ± 2 s.d. on CML versus normal $\log_2$ ratios. This candidate list was refined to include only those proteins corroborated by (1) log ratio changes of ±0.5 in murine Ba/F3 ± BCR–ABL1 MS data[12], (2) a complementary $CD34^+$ cell proteomics data set[39], and/or (3) all instruments within the current experimental data set. A parallel, manual inspection retained candidates if (1) log ratios were ≈ ±1.3, or (2) log ratios were lower and neither alternative instrument reported differential expression in the opposite direction; this manual selection step was blinded. Together, these filtering steps reduced the candidate list to 58 proteins (see Supplementary Table 1).

Formation of candidate network

The MetaCore implementation (13 June 2012) of Dijkstra's shortest path algorithm[7], a general purpose algorithm that identifies the shortest paths between 'seed' nodes of interest in a graph, was used to build a network around the 58 deregulated proteins (the graph used was the fully manually annotated MetaCore KB). Paths between seeds were limited to length = 2 and all shortest paths of the minimum length were retained in the resulting network. Topology statistics (Supplementary Table 2) were calculated using the igraph package in R.

Transcriptomic data analysis

Three bulk $CD34^+$ (1–3) and four primitive LSC (4–7) CML chronic phase versus normal data sets are discussed: (1) E-MTAB-2581: Affymetrix Human Gene 1.0 ST Array transcriptional

data from newly diagnosed CML chronic phase and normal progenitor cells ($CD34^+CD38^+$) from G-CSF-mobilized peripheral blood. mRNA was extracted using the RNeasy Mini Kit (Qiagen) and DNase I treated on columns using the RNase (RNase)-Free DNase Set (Qiagen). Affymetrix GeneChIP analysis was performed using 50 ng of RNA to the manufacturer's instructions; (2) Gene Expression Omnibus (GEO) accession GSE47927 (ref. 40); (3) GEO accession GSE5550 (ref. 41); (4) ArrayExpress accession E-MTAB-2581 (as described in (1) but for $CD34^+CD38^-$ cells); (5) GEO accession GSE47927 (ref. 40); (6) GEO accession GSE24739 (ref. 42); and (7) ArrayExpress accession E-MTAB-2508 (ref. 43).

CEL files were obtained directly from collaborators or public repositories. All transcriptional data were RMA normalized with the exception of GSE5550 for which only the VSN[44] normalized data were available. Log_2 scale expression values were analysed using limma[45] (*P* value correction by Benjamini–Hochberg[46]). In calculating the correlations and MI statistics (Fig. 1b, c and Extended Data Fig. 1), multiple probesets corresponding to single genes were median averaged.

Calculating correlations across multiple data sets

Ensembl's web services were used to map between (1) human and murine data sets via orthologues and (2) human transcriptomic and proteomic data sets via HUGO Gene Nomenclature Committee (HGNC) symbols. The Bioconductor package biomaRt (v.2.18.0) was used in R (v.3.0.1). Pearson's product moment correlation coefficients (r) were calculated across data sets for the 58 candidates and the 30 networked candidates (r_c and r_n respectively in Extended Data Fig. 1). Resulting r^2 values quantify the proportion of variability captured by the linear relationship and r^2_Δ is defined as r^2_n/r^2_c, that is, the ratio of r^2 for the 30 networked candidates to the r^2 for the 58 candidates. FDRs for the r^2_Δ observed were generated by considering 10,000 random samplings of each data set and counting the number of random samples meeting or exceeding the r^2_Δ and r^2_n statistics.

Calculation of MI

We calculated the MI values for 10,000 random subsets (of size 30) to generate a distribution of MI values that we would expect by chance (expression values binned between −3 and 3, bin width = 0.1; entropy package in R); these are the values summarized by the distributions plots in

Fig. 1c and Extended Data Fig. 1. The FDR values represent the proportion of random subsets that generate an MI greater than or equal to the MI for the network.

Calculation of proteomic/primitive transcriptional consistency FDR

Of the 30 proteins in Extended Data Fig. 1, 21 showed consistency of deregulation in at least three of the primitive transcriptional data sets and the bulk proteomic data set. To assess how likely it would be to observe such consistency by chance, the data were randomly permuted 10,000 times and permutations exhibiting similar deregulation consistency (that is, deregulation correspondence across four data sets) were recorded.

Topological analysis of p53/c-MYC in other MS data sets

Comparison of leukaemogenic PTK proteomic effects and three primary proteomics data sets, describing two types of breast cancer[11] (3 ductal carcinoma in situ and 4 invasive carcinoma breast cancer patient samples with matched normal samples), three types of prostate cancer[10] (24 non-aggressive, 16 aggressive, 25 metastatic prostate cancer patient samples and 10 normal samples) and cervical[9] cancer, were obtained directly from the authors. In each data set, deregulated proteins were identified using *z*-scores ± 2 when considering cancer versus normal $\log_2$ ratios and subjected to the same MetaCore network building process as described earlier. In the case of the cervical cancer data set no network could be found and the data were removed from the analysis. In addition, 50 sets of 58 random proteins were generated from the list of all proteins observed across the three MS data sets and subjected to the same network building process. The ratio d_{out}/d_{in} was calculated to quantify the bias of outgoing to incoming connections to/from p53 and c-MYC (Fig. 1d).

TKI-R/TKI-NR and aggressive/indolent CML transcription

The PB 'validation' set TKI-R/TKI-NR samples in GEO data set GSE14671 were integrated with a material-matched CML versus normal data set ($CD34^+CD38^+$ cell data from ArrayExpress E-MTAB-2581) using COMBAT[47] (Bioconductor package inSilicoDb v.1.10.1). All probeset-to-probeset mappings between the Affymetrix HG U133+2 and Affymetrix HuGe 1.0 ST chips (obtained via Bioconductor's biomaRt package) were retained and used by COMBAT. The TKI-R and TKI-NR samples were compared to the integrated normal data (using limma) to generate logFC values representing differential expression (see left two lanes in each panel of Fig. 5e). The pattern of differential expression in the TKI-R/TKI-NR versus normal comparisons was then

compared to that of CML versus normal, as calculated separately by limma in the material-matched data sets (E-MTAB-2581) (see right lane in each panel of Fig. 5e) to provide a baseline CML versus normal comparison. Transcriptional profiles for probe pairs corresponding to the 30 members of the candidate network were identified using HGNC symbols (data corresponding to TARDBP were removed due to the large number (66) of corresponding probe pairs). An FDR was calculated using 10,000 re-samplings of the merged data set to describe the likelihood of observing a correlation as high or greater by chance. The transcriptional data for the aggressive and indolent samples in ArrayExpress data set E-MIMR-17 were processed as described earlier, but integrated with the CML and normal $CD34^+$ data set GSE5550 using all probeset-to-probeset mappings between the Affymetrix U133a and Affymetrix HG Focus chips (again obtained using biomaRt).

Enrichment of MSigDB signatures

For the candidate network analysis (the results of which are shown in Supplementary Table 3), MSigDB signatures (C2: curated genesets) were accessed via the c2BroadSets object in GSVAdata v.1.0.0 using R. Enrichment scores were calculated using the hypergeometric distribution as implemented by dhyper(). Signatures corresponding to CML, c-MYC and p53 related biology were extracted using appropriate regular expressions on the signature name. In the RNA-seq analysis (Fig. 5b), the MSigDB signatures (C2: curated genesets) were identified as significantly differentially expressed in the TMM/VOOM-normalized RNA-seq data using GSVA[48] and limma[49]. Significant pathways were identified using an FDR = 0.05 threshold on corrected *P* values[46]. The p53, apoptosis, c-MYC and differentiation signatures were extracted using appropriate regular expressions on the signature name.

Enrichment of PANTHER pathways

The top 1,500 differentially expressed genes (as ranked by increasing *P* value, calculated by limma[49]) were identified comparing treated/untreated samples. These genes and their logFCs for each arm were uploaded to PANTHER (http://www.pantherdb.org/) and subjected to a Mann–Whitney *U*-test[50] to identify enrichment of PANTHER pathways. Extended Data Figure 6b shows enrichment results (without Bonferroni correction) including the hypothesized direction of pathway deregulation.

Cell counting and apoptosis assays

CML $CD34^+$ cells were seeded at $1–2 \times 10^6$ cells ml^{-1} before drug treatment and counted by trypan blue (Sigma-Aldrich) exclusion. Apoptosis was quantified by staining with annexin-V–APC and DAPI. In specified experiments, CML and human cord $CD34^+$ cells were labelled with CD34–APC and CD38–FITC or PerCP–Cy5.5 and sorted using the FACSAria (BD). The selected $CD34^+CD38^-$ cells were analysed 72 h after drug treatments. To measure the dose–effect relationship of each drug and its combination and to determine synergy, CIs were calculated using the Calcusyn software package (BioSoft). Except where documented, all results are expressed as a mean ± s.e.m.

CFC assay

$CD34^+$ cells were treated for 72 h at the indicated concentrations of RITA, CPI-203 and Das. Drug-treated cells (2,000 cells per plate) were washed and seeded in Methocult H4435 (STEMCELL Technologies). CML cells were transduced with lentivector constructs and sorted then washed and seeded into methylcellulose. Colonies were assessed 10–14 days after plating.

Western blotting

$CD34^+$ cells were lysed in RIPA buffer with inhibitors and western blots were performed as per standard protocols.

Immunofluorescence microscopy

CML $CD34^+$ cells were left untreated or treated with the indicated drugs for 24 h. Cells were harvested and spotted onto slides coated with poly-L-lysine and fixed with 3.7% (w/v) formaldehyde and permeabilized using a 0.25% (w/v) Triton-PBS solution for 15 min. Cells were blocked with 5% (w/v) BSA-PBS and stained with primary and secondary antibodies. Cells were concurrently stained with DAPI. Cells were imaged using a Zeiss Imager M1 AX10 fluorescence microscope (Carl Zeiss) and subjected to deconvolution (AxioVision software; Carl Zeiss) for image manipulation. Fluorescent signal was measured in three dimensions by Image Processing and Analysis in Java (Image J) program.

Tracking cell divisions

CFSE (Molecular Probes) and CD34 staining were performed and cell divisions were identified as described previously[3].

Lentivirus transduction

The pCMV-VSV-G and pCMV-HIV1 were provided by J. Rossi. The pLKO-GFP came from K. Kranc. The following optimized pLKO vectors were purchased from Open Biosystems and subcloned into the pLKO-GFP vector: (1) TRCN0000003380: MDM2 shRNA bacterial stock NM_002392.x-1495s1c1 (ref. 51); (2) TRCN0000355728: MDM2 shRNA bacterial stock NM_002392.3-1496s21c1; (3) TRCN0000174055: c-MYC shRNA bacterial stock NM_002467.2-1377s1c2 (ref. 52); (4) TRCN0000039642: c-MYC shRNA bacterial stock NM_002467.2-1377s1c1 (ref. 53).

Transduction of HeLa cell lines was performed at a MOI 1-10 with 70–95% of the cells expressing GFP after 48 h. For transduction, $CD34^+$ cells were cultured in medium supplemented with growth factors (IL-3 25 ng ml^{-1}, IL-6 10 ng ml^{-1}, Flt-3L 100 ng ml^{-1}, SCF 50 ng ml^{-1}, TPO 100 ng ml^{-1}) for 48 h, followed by two exposures to concentrated virus-containing supernatants (Multiplication of infection= 5) via spinoculation. Cells were harvested 48 h after second transduction and analysed or sorted for GFP positivity.

Transduced viable cells (assessed as annexin-V^-/$DAPI^-$ percentages multiplied by the absolute cell count) are presented as a percentage of CML $CD34^+$ cells transduced with scramble control.

Immunodeficient mouse engraftment

For the *ex vivo* drug studies CML (2×10^6 cells per mouse) or cord blood (2×10^5) $CD34^+$ cells were cultured with the indicated drugs (RITA 70 nM, CPI-203 1 μM and Das 150 nM). After 48 h cells were transplanted via tail vein into female 8–10-week-old sublethally irradiated (2.5 Gy) NOD.Cg-PrkdcscidIl2rgtm1Wjl/SzJ NSG mice (The Jackson Laboratory). Human cells were assessed by anti-human CD45 antibody analysed by flow cytometry. Specific cell subsets were detected using antibodies to human CD34, CD33, CD11b, CD14 and CD19 (mouse antibody table in Supplementary Table 8).

For the *in vivo* drug treated NSG experiments, CML (2×10^6 cells per mouse) $CD34^+$ cells were transplanted via tail vein into female 8–10-week-old sublethally irradiated (2.5 Gy) NSG mice (The Jackson Laboratory). After 4 weeks, mice were treated with RG7388 (75–100 mg kg^{-1}, oral gavage once daily), CPI-203 (6–7.5 mg kg^{-1}, intraperitoneally twice daily) or Nil (50 mg kg^{-1}, oral gavage once daily) for 3–4 weeks. Two CML samples were assessed

separately, each performed with 4–5 mice per drug arm per experiment. Results presented represent data from both experiments (each experiment normalized to vehicle). To quantify the frequency of BCR–ABL1^{+} cells within the engrafted human CD45^{+} cells, dual-fusion D-FISH was performed as previously described[3].

Transgenic mouse model

Inducible (tetracycline (TET)-based) DTG (SCLtTAxBCR-ABL) donor mice in a C57BL/6 (CD45.2) background were a gift from D. G. Tenen. B6.SJL-Ptprc^{a} Pepc^{b}/BoyJ (CD45.1) recipients (a mixture of female and males between 8–10 weeks old) were purchased from Charles River Laboratories.

Bone marrow transplantation and analysis of disease

Bone marrow cells of DTG mice (1×10^{6}) were injected into the tail veins of 10-week-old irradiated (2 doses of 4.25 Gy, 3 h apart) recipients. TET was continued for 2 weeks after radiation. Tail veins bleeds were performed weekly after TET removal and Gr1/Mac1 percentages (flow cytometry), white blood cells, neutrophils and haemoglobin (Hemovet) were monitored.

DTG *in vivo* drug treatment

Drugs were administered to DTG mice 5 weeks after transplantation, over a 4-week period. Nil 75 mg kg^{-1} once daily, CPI-0610 15 mg kg^{-1} twice daily and RG7112 50 mg kg^{-1} once daily, all by oral gavage. For no drug control mice were administered the vehicles at the same concentrations and volumes as used for the combination arm.

Flow analysis

Peripheral blood, bone marrow and spleen cells were stained using appropriate antibodies and analysed using a FACSCanto or FACSAria machine (BD Biosciences).

Husbandry

All experiments were performed in accordance with the local ethical review panel, the UK Home Office Animals Scientific Procedures Act, 1986, and UK Co-ordinating Committee on Cancer Research (UKCCCR) and National Cancer Research Institute (NCRI) guidelines.

Animals were kept in regulated facilities, monitored daily, and all experiments were carried out in compliance with UK Home Office guidelines. Mice were genotyped by Transnetyx.

RNA-sequencing

Sample preparation. CML $CD34^+$ cells were seeded in 48-well plates at $1–2 \times 10^6$ cells ml^{-1}, before drug treatment (RITA 50 nM, CPI-203 1 μM, Nil 5 μM) for 24 h. After treatment, RNA was extracted using RNeasy Plus Mini Kit (Qiagen).

Library generation. RNA-seq libraries were generated using TruSeq Stranded Total RNA (part no. 15031048 Rev. E October 2013) kits. Ribosomal depletion was performed on 1 μg of RNA using Ribo-Zero Gold before a heat fragmentation step aimed at producing libraries with an insert size between 120–200 bp. Complementary DNA was synthesized from the enriched and fragmented RNA using SuperScript II Reverse Transcriptase (Invitrogen) and random primers. The cDNA was converted into double-stranded DNA in the presence of dUTP to prevent subsequent amplification of the second strand. After 3′ adenylation and adaptor ligation, libraries were subjected to 15 cycles of PCR to produce RNA-seq libraries. Before sequencing, RNA-seq libraries were qualified and quantified via Caliper's LabChip GX (part no. 122000) instrument using the DNA High Sensitivity Reagent kit (product no. CLS760672). Quantification of libraries for clustering was performed using the KAPA Library Quantification Kits for Illumina sequencing platforms (kit code KK4824) in combination with Life Technologies QuantStudio 7 real-time PCR instrument. Libraries were finally pooled in equimolar ratios and sequenced on Illumina's NextSeq500 platform using 75 bp paired-end high-output runs.

Alignment and analysis. Sequencing reads were aligned to the genome (GRCh38/release 80 primary assembly as obtained via ftp.ensembl.org) using Subread (v.1.4.6-p3)[54]. RNA-SeQC was used to confirm adequate mapping quality and gene-level counts were calculated using Subread's featureCounts algorithm[55]. Count data for each arm were normalized independently by TMM[56] (as implemented in the Bioconductor package edgeR) and VOOM[57] (as implemented in the Bioconductor package limma). Genes with <3 cpm in three samples were removed from further analysis. Differential expression was identified using limma[49] (using Benjamini–Hochberg[46] correction).

Definition of synergy

Genes were described as loosely synergistic if (1) RITA, CPI-203 and combined treatment all induced deregulation in the same direction, (2) the deregulation in response to the combined treatment was significant ($q < 0.05$), and (3) the deregulation induced by the combined treatment was greater than both RITA and CPI-203 in isolation. A more extreme definition had an additional criteria of the log ratio of the observed and additive effect being >0.6 (corresponding to a 150% increase).

Statistical analysis

For *in vitro* experiments a minimum of 3 patient samples were chosen as a sample size to ensure adequate power. For all animal studies, each experiment was replicated twice in the laboratory with a minimum number of 5 mice per arm, unless indicated. NSG mice were excluded from analyses if they died of radiation poisoning (within 10 days of being irradiated, out of a 16-week procedure). For DTG mice, mice were excluded from analysis if leukaemic cells (CD45.2) failed to engraft host mice ($CD45.1^{+}$) and therefore would not develop leukaemia. This was determined 1 week before drug treatment. Patient samples were only excluded if clinical data identified patient sample as entering blast crisis. Pre-established criteria also included that if a sample data point deviated 2 standard deviations from the mean, it was to be excluded, but this was not applied to the data in main or extended data. Group allocation to mice was done as mice were either purchased and subsequently numbered or weaned, to remove any investigator bias. For both NSG and DTG mice studies, all mice were randomly assigned treatment groups, ensuring all animals were of equal health and leukaemic status, within normal variability. Mice were assessed at predetermined time points: NSG mice were assessed at 8 and 16 weeks and DTG mice were assessed 4 weeks after drug treatment, so there was minimized bias as to assessing outcome. All mice were cared for equally in an unbiased fashion by animal technicians and investigator. No blinding was done.

Unless indicated, data are presented as the mean ± s.e.m. and *P* values were calculated by two-tailed Student's *t*-test using GraphPad Prism software. Significant statistical differences ($*P < 0.05$, $**P < 0.01$, $***P < 0.001$) are indicated.

Code availability

All computer code was implemented in R and is available from the authors upon request.

39. Pizzatti, L., Sá, L. A., de Souza, J. M., Bisch, P. M. & Abdelhay, E. Altered protein profile in chronic myeloid leukemia chronic phase identified by a comparative proteomic study. *Biochim. Biophys. Acta* **1764,** 929–942 (2006).

40. Cramer-Morales, K. *et al.* Personalized synthetic lethality induced by targeting RAD52 in leukemias identified by gene mutation and expression profile. *Blood* **122,** 1293–1304 (2013).

41. Diaz-Blanco, E. *et al.* Molecular signature of $CD34^+$ hematopoietic stem and progenitor cells of patients with CML in chronic phase. *Leukemia* **21,** 494–504 (2007).

42. Affer, M. *et al.* Gene expression differences between enriched normal and chronic myelogenous leukemia quiescent stem/progenitor cells and correlations with biological abnormalities. *J. Oncol.* **2011,** 798592 (2011).

43. Graham, S. M., Vass, J. K., Holyoake, T. L. & Graham, G. J. Transcriptional analysis of quiescent and proliferating $CD34^+$ human hemopoietic cells from normal and chronic myeloid leukemia sources. *Stem Cells* **25,** 3111–3120 (2007).

44. Huber, W., von Heydebreck, A., Sültmann, H., Poustka, A. & Vingron, M. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* **18** (suppl. 1), S96–S104 (2002).

45. Smyth, G. K. in *Bioinformatics and Computational Biology Solutions using R and Bioconductor* (eds Gentleman, R., Dudoit, S., Irizarry, R. & Huber, W.) 397–420 (Springer, 2005).

46. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57,** 289–300 (1995).

47. Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8,** 118–127 (2007).

48. Hänzelmann, S., Castelo, R. & Guinney, J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* **14,** 7 (2013).

49. Smyth, G. K. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **3,** e3 (2004).

50. Clark, A. G. *et al.* Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios. *Science* **302,** 1960–1963 (2003).

51. Xu, X. L. *et al.* Retinoblastoma has properties of a cone precursor tumor and depends upon cone-specific MDM2 signaling. *Cell* **137,** 1018–1031 (2009).

52. Wolfer, A. *et al.* MYC regulation of a "poor-prognosis" metastatic cancer cell state. *Proc. Natl Acad. Sci. USA* **107,** 3698–3703 (2010).

53. Sheth, A. *et al.* Inhibition of human mitochondrial peptide deformylase causes apoptosis in c-myc-overexpressing hematopoietic cancers. *Cell Death Dis.* **5,** e1152 (2014).

54. Liao, Y., Smyth, G. K. & Shi, W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* **41,** e108 (2013).

55. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30,** 923–930 (2014).

56. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11,** R25 (2010).

57. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15,** R29 (2014).

Extended Data Figure 1 | BCR–ABL1 drives a proteomic signature mediated by p53 and c-MYC. **a**, **b**, Equivalent to Fig. 1b, c with additional information regarding the correlations calculated from the complete list of 58 candidate proteins (r_c) in addition to the correlations for the candidate network (r_n) and the background (r_0). Also shown is the gain in r^2 obtained for the candidate network as compared to the r^2 obtained for the candidate list as a whole (r^2_Δ). FDR calculated from 10,000 re-samplings. **c**, Expression changes of the network components (shown as bar plots) in the context of quiescent and primitive CML cells; data shown in each panel (left to right) are (1) $CD34^+$ protein $\log_2$ ratios (n=3 patient samples, n=2 normal samples); (2) $CD34^+Hst^{lo}Py^{lo}$ transcript logFC (ArrayExpress accession E-MTAB-2508); (3) $CD34^+Hst^{lo}Py^{lo}$ transcript logFC (GEO accession GSE24739); (4) $CD34^+CD38^-$ logFC (ArrayExpress accession E-MTAB-2581); and (5) $Lin^-CD34^+CD38^-CD90^+$ logFC (GEO accession GSE47927). Down-/upregulation is indicated by turquoise/red, respectively. Where multiple probesets were found for individual genes, the probeset corresponding to the maximal log ratio was selected. **d**, **e**,

Correlation of the candidate network in progenitor ($CD34^+$) CML cells: $CD34^+CD38^+$ progenitor (top); common myeloid progenitor $Lin^-CD34^+CD38^+CD123^+CD45RA^-$ (middle); and $CD34^+$ cells (bottom). As in **a**, **b**, correlations for the background (r_0), candidate list (58 proteins, r_c) and candidate network (Fig. 1a, r_n) are shown. Also shown is the gain in r^2 obtained for the candidate network as compared to the r^2 obtained for the candidate list as a whole (r^2_Δ). FDR calculated from 10,000 re-samplings; MI statistics corresponding to FDRs<0.05 are coloured red, FDRs<0.10 are coloured grey. **f**, A Venn diagram showing the overlap in protein identification of the three MS instruments: ABSciex Q-STAR Elite (Elite), Thermo LTQ Orbitrap Velos (Orbi) and ABSciex TripleTOF 5600 (5600).

Extended Data Figure 2 | Validation of network candidates. **a**, HDM2 and c-MYC knockdown using shRNA constructs. Western blots of c-MYC, HDM2, p53 and Hsp90 in HeLa cells transduced with lentiviral constructs specific for either c-MYC (2 constructs), HDM2 (2 constructs) or scrambled control (1 construct). KD, knockdown. **b**–**d**, CML $CD34^+$ cells were transduced with either lentiviral (GFP) shRNA constructs to HDM2 (constructs 1, 2), c-MYC (constructs 1, 2) or scramble control (1 construct). **b**, Transduced viable GFP^+ cells (assessed as annexin-V^-/$DAPI^-$/GFP^+ percentages multiplied by the absolute cell count) are presented as a percentage of CML $CD34^+$ cells transduced with scramble control (n = 3 patient samples). **c**, Early apoptosis levels (assessed as annexin-V^+/$DAPI^-$/GFP^+) after transduction of CML $CD34^+$ cells (n = 3 patient samples) as described in **b**. Statistical significance was calculated by a two-tailed Student's t-test and error bars represent the s.e.m.

Extended Data Figure 3 | RITA and CPI-203 synergize to eliminate CML $CD34^+$ cells. a, **b**, Western blots of CML $CD34^+$ cells untreated or treated with 50 nM RITA; 1 μM CPI-203; or the combination of 50 nM RITA and 1 μM CPI-203 or 150 nM Das for 8 h (**a**) and 48 h (**b**). **c**, p53 (red, nucleus in blue) 24 h after treatment in CML $CD34^+$ cells. **d**, RITA, CPI-203 and combination drug treatment eliminates $CD34^+$ CML cells through mechanisms probably dependent on apoptosis; after 72 h of drug treatment apoptosis levels were assessed (annexin V/DAPI) using flow cytometry techniques. **e**, RITA or Nut cannot induce death of K562 cells that lack p53. K562 cells were treated with either 50 μM RITA or 10 μM Nut and after 72 h of drug treatment, apoptosis levels were assessed (annexin V/DAPI) using flow cytometry techniques.

Extended Data Figure 4 | RITA and CPI-203 synergize to eliminate CML cells a, $CD34^+$ CML cells were treated with Nut and CPI-203 for 72 h with apoptosis levels assessed (annexin V/DAPI) using flow cytometry techniques. **b**, Treatment of $CD34^+$ CML cells with Nut results in the elimination of early and late progenitor cells as assessed by the functional colony-forming capacity of drug-treated CML cells. **c**, Sequential drug treatments (n = 3 patient samples; drug one for 24 h, then both for 48 h). **d**, Sequential knockdown treatments (n = 3 patient samples; knockdown one for 24 h, then both for 48 h), mean ± s.e.m. (P values: two-tailed Student's t-test; $*P < 0.05$, $**P < 0.01$, $***P < 0.001$). **e**, CML $CD34^+$ primary samples were pre-treated or not with imatinib mesylate (1 μM) for 8 h followed by RITA (50 nM), CPI-203 (1 μM) or the combined treatment (RITA plus CPI-203) for 72 h (right three columns). Cell counts were obtained using trypan blue exclusion.

Extended Data Figure 5 | RITA and CPI-203 selectively eliminate LSCs. **a**, **b**, Viable cell counts (n = 3 patient samples) (**a**); apoptosis in normal $CD34^+$ cells (n = 3 patient samples) in response to RITA and/or CPI-203 (**b**). **c**, Gated CML $CD34^+CD38^-$ cells 72 h after treatment (n = 4 patient samples). **d**, *Ex vivo* protocol for CML/cord blood $CD34^+$ cells in NSG mice (n = 5 mice per arm). **e**, **f**, Targeting p53 and c-MYC in CML eliminates NSG repopulating leukaemic stem cells. CML $CD34^+$ cells were treated with RITA (70 nM) and/or CPI-203 (1 μM) or Das (150 nM) for 48 h and recovered cells were injected intravenously into 8–12-week-old, sublethally irradiated (2.5 Gy) NSG mice (2–4 mice per arm). **e**, Percentage of human $CD45^+$ cell levels in peripheral blood (PB) at 8, 12 and 16 weeks. **f**, Percentages of human $CD45^+$, $CD34^+$, $CD33^+$, $CD11b^+$, $CD19^+$ and $CD14^+$ cells in the bone marrow at 16 weeks. **g**, CML bone marrow analyses of CD33, CD11b, CD19 and CD14 from a CML sample determined to engraft both BCR–ABL-positive and -negative cells. **h**, D-FISH analyses of bone marrow human engraftment studies shown in **g** performed twice (2 patient samples) with a minimum of n = 6 mice per arm; mean ± s.e.m. (P values: two-tailed Student's t-test; $*P < 0.05$, $**P < 0.01$, $***P < 0.001$).

Extended Data Figure 6 | Mechanism of LSC elimination and clinical scope. **a**, Enrichment of p53 (top); apoptosis (second from top); c-MYC (second from bottom); and differentiation MSigDB signatures (bottom) in the four treatment arms (n=3 CML patient samples per arm) (columns named as per **b**). Equivalent to Fig. 5b, but with named MSigDB signatures. **b**, Enrichment of PANTHER pathways in the four treatment arms. Pathway enrichment calculated

from the top 1,500 genes, as ranked by increasing *P* value (top); only those genes exhibiting an absolute FC of >0.5 in each arm (bottom left); only those genes exhibiting a *P* value of >0.05 in each arm (bottom right). **c**, Assessing molecular synergy of the combined RITA plus CPI-203 treatment, as compared to the individual RITA and CPI-203 arms of the RNA-seq experiments in the three *in silico* functional signatures: p53/apoptosis (left); c-MYC (middle); and differentiation (right). Mean expression is shown as a solid line.

Extended Data Figure 7 | Mechanism of LSC elimination and clinical scope continued. **a**, Gene expression patterns (logFC, n=3 patient samples per arm) shown for the members of the three broad signatures identified *in silico*: p53/apoptosis (left); c-MYC (middle); and differentiation (right) (**q* < 0.05); data are ordered by increasing logFC in response to Combo treatment, from downregulation at the top to upregulation at the bottom. Corresponding expression data are provided in Supplementary Tables 5–7. **b**, Differential expression of CD34 and CD133 (markers of stemness) in the four arms of the RNA-seq experiment. **c**, Apoptosis levels assessed (annexin V^+/DAPI) using flow cytometry on a TKI-NR $CD34^+$ sample after 72 h treatment with RITA and CPI-203 as indicated.

Extended Data Figure 8 | RG7112 and CPI-0610 as a combination decrease BCR–ABL$^+$ cells. **a**, **b**, DTG mice *in vivo* treatment (**a**): neutrophils normalized to control (dotted line) (**b**). **c**, Bone marrow cells stained for CD45.1/2. Drug treatment arms (minimum of *n* = 7 mice) mean ± s.e.m. (*P* values: two-tailed Student's *t*-test; **P* < 0.05, ***P* < 0.01, ****P* < 0.001).