The EMBO Journal (2012) 31, 3704–3717 | © 2012 European Molecular Biology Organization | Some Rights Reserved 0261-4189/12
www.embojournal.org

THE EMBO JOURNAL

EMBO open

A novel molecular mechanism involved in multiple myeloma development revealed by targeting MafB to haematopoietic progenitors

Carolina Vicente-Dueñas[1,2], Isabel Romero-Camarero[1,2], Inés González-Herrero[1,2], Esther Alonso-Escudero[1,2], Fernando Abollo-Jiménez[1,2], Xiaoyu Jiang[3], Norma C Gutierrez[2,4], Alberto Orfao[2,5], Nieves Marín[6], Luisa María Villar[6], Mª Carmen Fernández Criado[7], Belén Pintado[8], Teresa Flores[2,9], Diego Alonso-López[10], Javier De Las Rivas[10,11], Rafael Jiménez[2,12], Francisco Javier García Criado[2,13], María Begoña García Cenador[2,13], Izidore S Lossos[3,14], César Cobaleda[15,*] and Isidro Sánchez-García[1,2,*]

[1]Experimental Therapeutics and Translational Oncology Program, Instituto de Biología Molecular y Celular del Cáncer, CSIC/Universidad de Salamanca, Salamanca, Spain, [2]Institute of Biomedical Research of Salamanca (IBSAL), Salamanca, Spain, [3]Division of Hematology-Oncology, Sylvester Comprehensive Cancer Center, University of Miami, Miami, FL, USA, [4]Servicio de Hematología. Hospital Universitario de Salamanca, Salamanca, Spain, [5]Servicio de Citometría and Departamento de Medicina, Universidad de Salamanca, Salamanca, Spain, [6]Servicio de Inmunología, Hospital Ramón y Cajal, Madrid, Spain, [7]Facultad de Medicina, Universidad Autónoma, Madrid, Spain, [8]Genetically Engineered Mouse Facility, CNB-CSIC, Madrid, Spain, [9]Departamento de Anatomía Patológica, Universidad de Salamanca, Salamanca, Spain, [10]Bioinformatics Unit, Cancer Research Center (CSIC-USAL), Salamanca, Spain, [11]Bioinformatics and Functional Genomics Research Group, Cancer Research Center (CSIC-USAL), Salamanca, Spain, [12]Departamento de Fisiología y Farmacología, Universidad de Salamanca, Edificio Departamental, Salamanca, Spain, [13]Departamento de Cirugía, Universidad de Salamanca, Salamanca, Spain, [14]Department of Molecular and Cellular Pharmacology, University of Miami, Miami, FL, USA and [15]Centro de Biología Molecular Severo Ochoa, CSIC/Universidad Autónoma de Madrid, Madrid, Spain

Understanding the cellular origin of cancer can help to improve disease prevention and therapeutics. Human plasma cell neoplasias are thought to develop from either differentiated B cells or plasma cells. However, when the expression of *Maf* oncogenes (associated to human plasma cell neoplasias) is targeted to mouse B cells, the resulting animals fail to reproduce the human disease. Here, to explore early cellular changes that might take place in the development of plasma cell neoplasias, we engineered transgenic mice to express *MafB* in haematopoietic stem/progenitor cells (HS/PCs). Unexpectedly, we show that plasma cell neoplasias arise in the MafB-transgenic mice. Beyond their clinical resemblance to human disease, these neoplasias highly express genes that are known to be upregulated in human multiple myeloma. Moreover, gene expression profiling revealed that MafB-expressing HS/PCs were more similar to B cells and tumour plasma cells than to any other subset, including wild-type HS/PCs. Consistent with this, genome-scale DNA methylation profiling revealed that MafB imposes an epigenetic program in HS/PCs, and that this program is preserved in mature B cells of MafB-transgenic mice, demonstrating a novel molecular mechanism involved in tumour initiation. Our findings suggest that, mechanistically, the haematopoietic progenitor population can be the target for transformation in MafB-associated plasma cell neoplasias.

The EMBO Journal (2012) **31,** 3704–3717. doi:10.1038/emboj.2012.227; Published online 17 August 2012
Subject Categories: genome stability & dynamics; molecular biology of disease
Keywords: cancer therapy; MafB; multiple myeloma mouse model; oncogenes; reprogramming stem cells

Introduction

The identification of the cells of origin from which cancer initially arises is of great importance, both for our understanding of the basic biology of tumours and for the translation of this knowledge to the prevention, treatment, and precise prognosis of the human disease (Visvader, 2011). Traditionally, the identity of the cancer cell-of-origin was extrapolated from the histological characterization of tumours, and assimilated to the most analogous physiological cellular type. However, several transcriptome studies have shown that the molecular characteristics of tumoral cells do not correspond, in many cases, to what they seem to be according to their appearance under the microscope (Lim *et al*, 2009). For this reason, extrapolating the identity of the cancer cell-of-origin without appropriate functional lineage tracing analyses can lead us to the wrong conclusions (Molyneux *et al*, 2010).

Multiple myeloma (MM) is a malignancy characterized by the abnormal expansion of the terminally differentiated cells of the B-cell lineage, the plasma cells (Jaffe *et al*, 2001), and the nature of its cell-of-origin is still controversial. In fact, the term tumour/cancer stem cell was first coined nearly 40 years ago to highlight the observation that only a minority of MM cells were capable of clonogenic growth (Hamburger and Salmon, 1977). Two different functional approaches have been used to directly investigate the identity of the cell of origin of MM. One approach relies on the use of genetically

*Corresponding authors. C Cobaleda, Centro de Biología Molecular Severo Ochoa, CSIC/Universidad Autónoma de Madrid, c/Nicolás Cabrera, no 1, Campus de Cantoblanco, Madrid 28049, Spain. Tel.: +34 911964692; Fax: +34 911964420;
E-mail: ccobaleda@cbm.uam.es or I Sánchez-García, Experimental Therapeutics and Translational Oncology Program, Instituto de Biología Molecular y Celular del Cáncer, CSIC/Universidad de Salamanca, Campus M. de Unamuno s/n, Salamanca 37007, Spain. Tel.: +34 923238403; Fax: +34 923294813; E-mail: isg@usal.es

Received: 17 May 2012; accepted: 20 July 2012; published online: 17 August 2012

engineered mouse models to induce the ectopic expression of oncogenes associated to human plasma cell neoplasias in different stages of B-cell development. Therefore, these studies utilized B cell-specific promoters, like the IgH promoter or the Eμ enhancer, to overexpress, in the B-cell lineage, plasma cell neoplasia-associated oncogenes, like c-Maf (Morito *et al*, 2011). However, what these mice develop are B-cell lymphomas with some clinical features that resembled those of MM (Morito *et al*, 2011). Thus, this approach of targeting oncogenes associated to human plasma cell neoplasias to mouse B cells has failed to fully recapitulate the human MM disease. The other approach is to investigate the tumorigenic potential of MM cells (either derived from patients or from myeloma-prone mouse strains) by *in-vivo* transplantation assays. Using this approach, it has been demonstrated that memory B cells, rather than mature plasma cells, are able to produce symptomatic disease in immunodeficient mice (Matsui *et al*, 2004, 2008). These *in-vivo* transplantation-based approaches are designed to identify the tumour-propagating cells, but not the cell-of-origin; hence, these studies cannot exclude that progenitor cells (PCs) could also serve as the cells of origin for MM *in vivo*. In summary, due to a lack of mouse models that enable strict lineage targeting in the haematopoietic stem (HS)/PC population, it has not been comprehensively determined previously whether these cells can serve as targets for transformation in MM. The present study aims to reveal the role of PCs in MM initiation, using *MafB* as a model of an oncogene associated to human plasma cell neoplasias.

MafB is a member of the Maf family proteins, which are basic-leucine zipper transcription factors with important functions both in early tissue specification and in terminal differentiation (Eychene *et al*, 2008). The expression of Maf proteins is tightly regulated in a spatio-temporal manner during development (Eychene *et al*, 2008). MafB is an inducer of monocytic differentiation that is expressed in myeloid cells and precursors throughout haematopoietic differentiation (Kelly *et al*, 2000). Expression of MafB in erythroblasts inhibits erythroid differentiation (Sieweke *et al*, 1996). However, MafB expression induces the monocyte commitment of human $CD34^+$ stem/progenitor cells (Gemelli *et al*, 2006) and selectively restricts myeloid commitment divisions at the haematopoietic stem cell (HSC) in the mouse, hence contributing to the maintenance of a balanced lineage potential in the HSCs (Sarrazin *et al*, 2009). Maf proteins have been directly implicated in carcinogenesis, both in cell culture systems and in human cancers (Eychene *et al*, 2008). Among the different MAF proteins, MAFA and c-MAF display the strongest oncogenic activity, whereas MAFB is less effective in transforming cells (Nishizawa *et al*, 2003; Pouponnot *et al*, 2006). Translocations affecting either *c-MAF* (16q23) or *MAFB (20q12)* are present in 8–10% of the cases of MM (Mitsiades *et al*, 2004; Tosi *et al*, 2006; Hideshima *et al*, 2007). Even when it is not involved in translocations, *c-MAF* overexpression has been found in 50% of MM bone marrow (BM) samples and several human MM cell lines, suggesting an essential role for this MAF family in the pathobiology of MM (Hurt *et al*, 2004). Recently, it has been suggested that a threshold level of MAF expression might be required for transformation, as only mice carrying a high copy number of the MAF transgene in the T-cell lymphoid compartment develop T-cell lymphoma (Morito *et al*, 2006) or, in the B-cell lymphoid compartment, B-cell lymphomas (Morito *et al*, 2011).

In this study, we have explored the early cellular changes that might occur in plasma cell neoplasias by engineering transgenic mice to express *MafB* in HS/PCs. Unexpectedly, we show that plasma cell neoplasias arise in the MafB transgenic mice. Besides their clinical resemblance to human disease, these plasma cell neoplasias highly express genes that are known to be upregulated in human MM. Moreover, gene expression profiling revealed that MafB-expressing HS/PCs were more similar to B cells and tumour plasma cells than to any other subset, including wild-type HS/PCs. Consistent with this, genome-scale DNA methylation profiling revealed that MafB imposes an epigenetic program in HS/PCs, and that this program is preserved in mature B cells of Sca1-MafB mice, therefore showing that *MafB* can act as a reprogramming factor to reset the genome of stem/precursors cells to a terminally differentiated tumour state. Overall, our findings suggest that a haematopoietic progenitor population can be a target for transformation in MafB-associated plasma cell neoplasias.

Results

Detection of chromosomal translocations by fluorescence in-situ hybridization in BM $CD34^+$ cells of MM patients

Chromosomal translocations involving the immunoglobulin heavy chain (IGH) gene are detected in 50–60% of MM patients, using fluorescence *in-situ* hybridization (FISH). These chromosomal rearrangements have been traditionally used to identify tumour plasma cells. However, their presence has not been investigated in primitive HS cells. Therefore, in our aim of identifying the cell of origin for MM, we initially searched for the presence of chromosomal translocations in the haematopoietic/progenitor stem cells of MM patients. To this aim, BM $CD34^+$ cells (a marker for human haematopoietic progenitor and stem cells) and $CD38^+CD138^+$ cells (markers for plasma cells) were isolated from 15 MM patients. FISH analysis of the $CD38^+CD138^+$ cells showed the presence of *IGH* rearrangements in eight of the cases. However, the $CD34^+$ cells did not seem to show any of the translocations detected in $CD38^+CD138^+$ cells (Supplementary Figure 1; Supplementary Table I), although these aberrations would be difficult to detect if the frequency of these putative stem cells harbouring the translocation was low. These results, at face value, would seem to suggest that, either the MM cell-of-origin where oncogene activation takes place as a result of a chromosomal rearrangement is not a stem/progenitor cell or that, being a stem/progenitor cell, the oncogene might downregulate the stem cell markers (Mak *et al*, 2012). However, until now, all the experiments targeting the expression of human plasma cell neoplasia-associated cMaf oncogenes to the mouse B-cell compartment have failed to reproduce the human disease in mice. From other types of tumours, like for example acute leukaemias (Cobaleda and Sanchez-Garcia, 2009), it is known that the functional impact of the oncogenic translocations can manifest in the form of cellular types whose markers of differentiation place them either downstream or upstream of the point of origin of the translocation. Therefore, it is potentially possible that, in human patients, the occurrence of MM-associated oncogenic

alterations might happen in the HS/progenitor compartment, and this cell-of-origin adopts/acquires afterwards a tumoral plasma cell fate as a consequence of the oncogene's activity. This last possibility would reconcile the aforementioned findings derived from both human and mouse models. Therefore, in order to explore early cellular changes that could/might occur in plasma cell neoplasias, we engineered transgenic mice to express *MafB* in HS/PCs.

In-vivo ectopic expression of MafB in the HS/progenitor compartment

Our initial strategy was therefore to determine whether enforced expression of the MafB transcription factor was sufficient to reprogram HS/PCs into tumour plasma cells. Thus, in order to verify the feasibility of our experimental approach, we initially examined the expression of the endogenous MafB in normal HS/PCs ($Sca1^+Lin^-$) cells purified from the BM of control mice. *MafB* mRNA levels were undetectable by RT–PCR analysis in the $Sca1^+Lin^-$ compartment of control mice (Supplementary Figure 2A). Consequently, our strategy is an ideal *in-vivo* model to study the consequences of ectopic *MafB* expression in HS/PC $Sca1^+$ cells.

We next generated and characterized transgenic mice engineered to express the *MafB* cDNA under the control of the Sca1 promoter, in order to determine the effect of ectopic expression of *MafB* in HS/PCs (Figure 1). The Sca1-based system ensures the expression of MafB into the HS/PCs compartment (Miles *et al*, 1997; Perez-Caro *et al*, 2009; Vicente-Duenas *et al*, 2012). Insertion of the *MafB* cDNA under the control of the mouse Ly-6E.1 promoter (Miles *et al*, 1997) yielded the plasmid Sca1-MafB (Figure 1A), which was used to drive Sca-1-directed expression of MafB in C57BL/$6\times$CBA mice (Figure 1B). Two founders were obtained for the Sca1-MafB transgene, and Southern blot comparison of the endogenous and transgenic MafB hybridization signals indicated transgene copy numbers ranging from $\sim$2 to 4 (Figure 1B). Both independent Sca1-MafB founder lines had normal gestation and were viable and were used to examine the phenotype further (Table I). Since c-kit is known from previous studies to be downregulated in leukaemia stem cells (Blair and Sutherland, 2000; Neering *et al*, 2007), our functional definition of stem cell in this study does not include c-kit as a surface marker.

In agreement with the known capacity of Sca1 regulatory elements to drive and restrict transgenic expression to $Sca1^+$ cells, Sca1-MafB transgene expression was detected in purified $Sca1^+Lin^-$ cells as determined by RT–PCR (Figure 1C; Supplementary Figure 2B and C). On the contrary, Sca1-MafB transgene expression was not detected in purified plasma cells ($B220^{low}CD138^{hi}FSC^{hi}SSC^{hi}$) (Figure 1C; Supplementary Figure 2B). Together, these results documented limited increased levels of ectopic MafB expression only in $Sca1^+$ cells, enabling an *in-vivo* analysis of the functional impact of enforced MafB expression in the stem/progenitor cell compartment. To this end, Sca1-MafB mice were assessed using flow cytometric and histological analysis.

Reprogramming of HS/PC cells to terminally differentiated plasma cells in the Sca1-MafB mice

Since MafB is an inducer of monocytic differentiation and blocks other haematopoietic lineages (Sieweke *et al*, 1996; Kelly *et al*, 2000), all the main haematopoietic compartments were initially studied using flow cytometry in young Sca1-MafB mice. At 8 weeks of age, no major abnormalities could be detected neither in the myeloid nor in the T-lymphoid compartments (as determined by Gr1, Mac1 or CD4, CD8 stainings) in the peripheral blood (PB), BM, spleen, and thymus (Supplementary Figures 3–6). Also at 8 weeks, B-lymphoid and plasma cell compartments in the PB were comparable to those from wild-type control mice (as analysed with B220, CD138, and IgM stainings). However, in the BM, alterations of B-cell development could already be detected at this time (8 weeks) in the form of changes in the proportions of the different $B220^+$ cell compartments. By 12 months of age (Supplementary Figure 7), the IgM^-B220^{dull} B-cell progenitor population had practically disappeared from the BM of Sca1-MafB mice, and this was correlated with a decreased percentage of B cells in the PB, BM, and spleen, indicating a clear defect in B lymphopoiesis. The cells of human plasma cell myeloma (the human haematopoietic cancer most frequently associated with MafB overexpression) typically present high levels of syndecan-1 (CD138) and low levels or absence of expression of the pan-B cell antigen (CD19 in humans or B220 in mice). In the BM and spleen of Sca1-MafB mice, accumulations of $B220^{low}CD138^{hi}FSC^{hi}SSC^{hi}$ plasma cells could be clearly seen by 12 months of age (Supplementary Figure 7B and C). In summary, the analysis of Sca1-MafB mice by flow cytometry showed a normal development of all the haematopoietic compartments that was paralleled by the slow development of a plasma cell accumulation in aging mice.

Reprogrammed plasma cells behave as tumour plasma cells

Sca1-MafB mice exhibited an overall shortened lifespan when compared with WT littermates (Figure 1D). The morphologic analysis of the BM of Sca1-MafB showed the accumulation of plasma cells (Figure 2A). In contrast to normal or reactive plasma cells, which usually occur in small clusters of five or six cells around marrow arterioles, plasma cell accumulations in Sca1-MafB mice frequently occurred in larger loci, nodules or sheets, in such a way that a significant BM volume was comprised of plasma cells and plasmablasts. This strongly supports the malignant nature of the plasma cells. As shown in Figure 2A, reprogrammed plasma cells varied from mature forms indistinguishable from normal plasma cells to immature, pleomorphic or anaplastic plasma cells. The mature plasma cells were usually oval, with a round eccentric nucleus with the typical 'spoke wheel' chromatic structure, with abundant basophilic cytoplasm and a marked perinuclear hof. In contrast, immature forms had dispersed nuclear chromatin, a high nuclear/cytoplasmic ratio, and prominent nucleoli (plasmablasts). Because nuclear immaturity and pleomorphism rarely occur in reactive plasma cells, they are reliable indicators of neoplastic plasmacytosis.

Upon necropsy of the sacrificed mice, macroscopic analysis of the organs showed obvious splenomegaly and renal pathology (Supplementary Figure 8). Histological examination of the Sca1-MafB kidneys revealed aggregates of eosinophilic material (see below) in the lumen of renal tubules, responsible for the tubular and glomerular damage (Supplementary Figure 9A). The histological analysis also revealed infiltration of either plasma cells or their

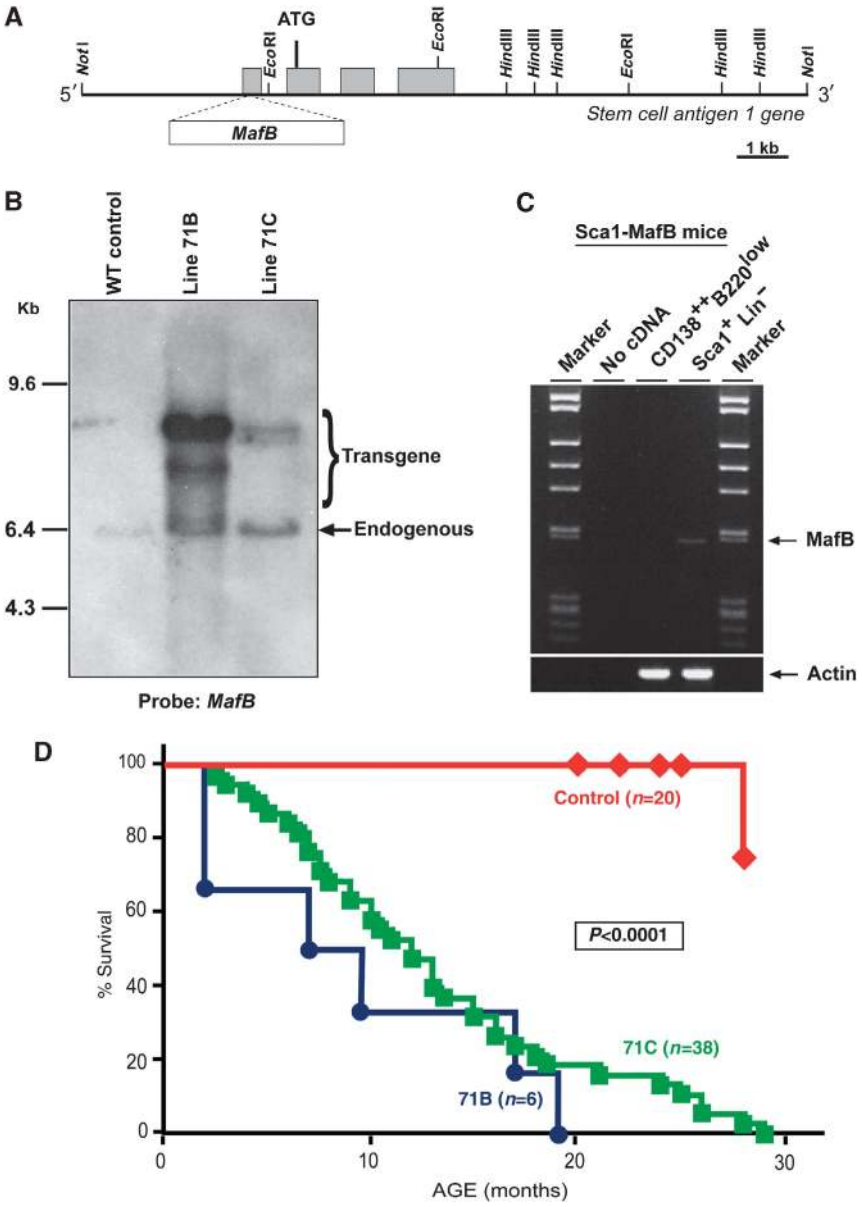


Figure 1 Transgenic expression of MafB in the mouse Sca1 lineage. (**A**) Schematic representation of the genomic structure of the mouse *Sca1* locus and the Sca1-MafB transgenic vector used in this study. *Not*I sites used to excise the transgene fragments and *Eco*RI and *Hin*dIII sites used to examine Southern blots are indicated. (**B**) Identification of the founder transgenic mice of the two lines used in this study by Southern analysis of tail snip DNA after *Hin*dIII digestion. *MafB* cDNA was used for the detection of the transgene. Sca1-MafB and the endogenous MafB are indicated. (**C**) *MafB* expression in stem and plasma cells of Sca1-MafB mice. RT–PCR analysis revealed *MafB* expression in stem ($Sca1^{+}Lin^{-}$) cells but not in plasma ($B220^{low}CD138^{hi}FSC^{hi}SSC^{hi}$) cells purified from the BM of Sca1-MafB mice (line 71C). Actin was used to check cDNA integrity and loading (bp, base pairs). The correct identity of the PCR product was analysed by hybridization using as a probe the *MafB* cDNA (not shown). (**D**) Kaplan–Meier survival plots of Sca1-MafB mice (lines 71B or 71C). The total number of mice analysed in each group is indicated. Statistical analysis of differences in survival was performed using the log-rank (Mantel-Cox) test, and the corresponding *P*-value is given.

Table I Incidence and age of plasma cell neoplasia onset in Sca1-MafB mice

Transgenic line	Mice autopsied[a]	Mice with tumour (%)[b]	Age at tumour onset (months)	Tumour type (%)	Mice with plasma cell accumulation (%)	Mice with amyloidosis (%)	Osteolitic lesions (%) ($n=9$)	Oligoclonal IgG increase (%) ($n=8$)
71B	6	3 (50)	15.2 ± 4.09	Plasma cell neoplasia (100)	3 (50)	ND	1/1 (100)	ND
71C	38	17 (44.7)	15.2 ± 7.5	Plasma cell neoplasia (100)	17 (44.7)	18/38 (47.37)	4/8 (50)	7/8 (87)

[a]Number of mice during or after the period of cancer.
[b]Number of mice killed with cancer and percentage of tumour incidence; n, number of mice studied; ND, not determined.

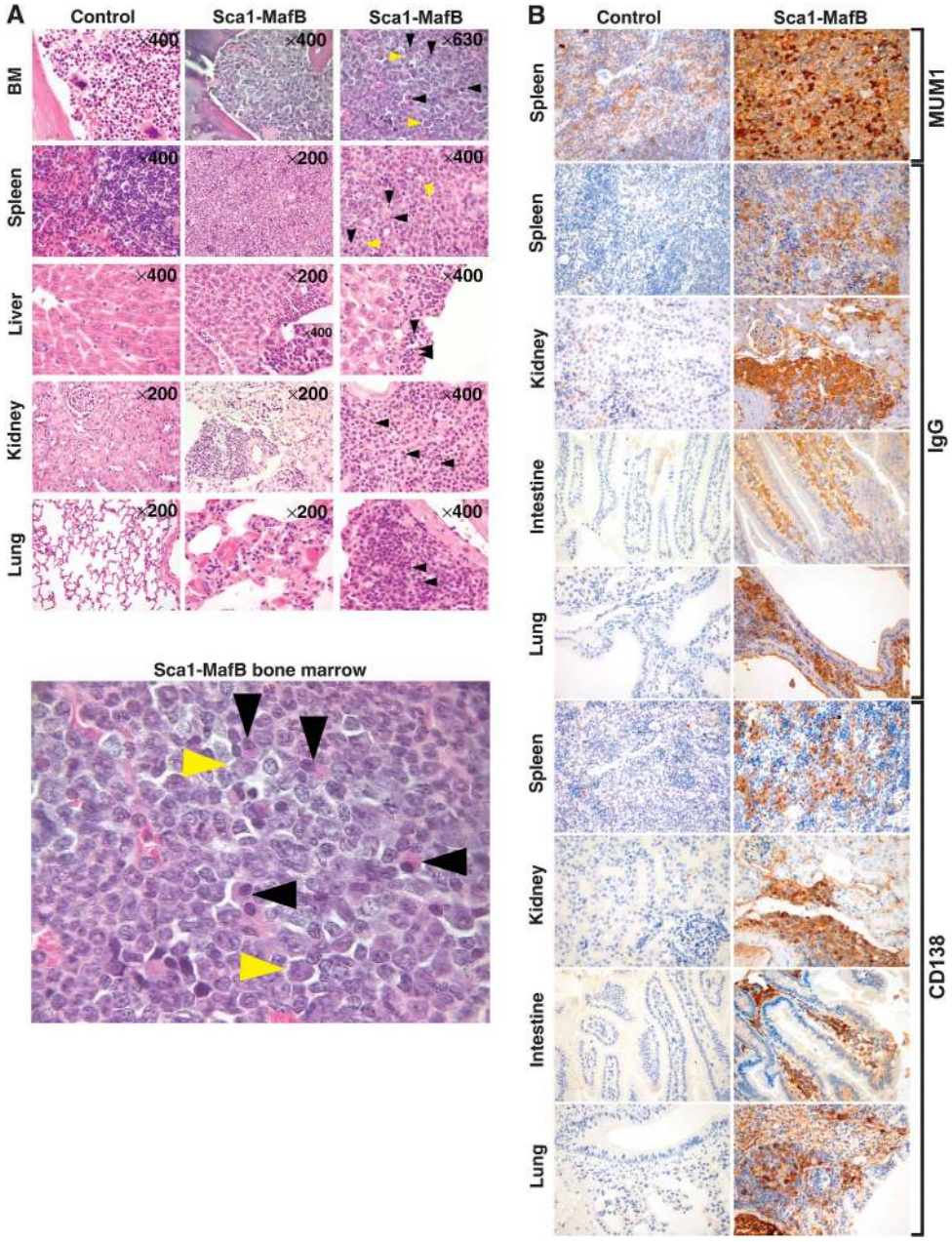


Figure 2 Reprogramming of Sca1 cells to terminally differentiated plasma cells. (**A**) Representative histological appearance of haematopoietic (BM, spleen) and non-haematopoietic (liver, kidney, and lung) tissues of diseased Sca1-MafB mice versus WT control mice after haematoxylin-eosin staining. A total number of 44 Sca1-MafB mice were analysed, out of which 20 showed accumulations of plasma cells. Bottom panel shows an enlargement of the BM of Sca1-MafB mice, where the infiltrating plasma cells and plasmablasts characteristic of multiple myeloma can be clearly recognized. Examples of plasma cells and plasmablasts are indicated by black and yellow arrows, respectively. (**B**) Representative immunohistochemical analysis of the tissues of diseased Sca1-MafB mice versus age-matched WT controls. To prove the identity of the infiltrated cells beyond their plasma cell morphology, immunohistochemical analysis was performed in four different transgenic mice using the plasma cell-specific markers Mum1/IRF4 or CD138 in different organs, as indicated. The presence of plasma cells was confirmed in all the organs analysed. Furthermore, by staining with an anti-IgG antibody the infiltration of plasma cells could be detected in spleen, kidney, intestine, and lung. The image magnification is $\times 400$ in all cases.

plasmablast precursors in haematopoietic and non-haematopoietic organs, like BM, spleen, liver, kidney, and lungs (Figure 2A). To prove the identity of the infiltrated cells beyond their plasma cell morphology, immunohistochemical analysis was performed using the plasma cell-specific markers Mum1/IRF4 and CD138 (Figure 2B). The presence of $IRF4^+$ or $CD138^+$ cells was confirmed in spleen, kidney, intestine, and lung. Also staining with an anti-IgG antibody revealed the infiltration of plasma cells in all these organs (Figure 2B). These infiltrations were associated with an increased number of plasma cells in the BM of Sca1-MafB mice compared with that of wild-type control mice (Figure 2A; Supplementary Figure 7). Thus, the macroscopic and histological findings define the malignant nature of the reprogrammed plasma cells and further support the plasma cell accumulation that was already detected by flow cytometry, especially considering

that it is well established that flow cytometry underestimates the number of plasma cells due to their fragility and high adhesivity (Orfao *et al*, 1994; Ng *et al*, 2006).

Lytic bone lesions and tumoral masses of plasma cells are one of the most characteristic hallmarks of malignant plasma cell behaviour. The most common sites of appearance of these lesions are the BM areas where the haematopoiesis is most active including, in order of frequency, the vertebrae, ribs, skull, pelvis, femur, clavicle, and scapula (Jaffe *et al*, 2001). A detailed X-ray analysis of Sca1-MafB diseased (between 9 and 16 months old) mice ($n=9$) showed the presence of clear osteolytic lesions in 55% of them (5/9) (Figure 3A). These osteolytic lesions were produced as a consequence of the accumulation of plasma cells around the bone lytic area, as evidenced by histological analysis in the Sca1-MafB mice (Figure 3B). Even more, in some instances the extensive skel*etal* destruction by neoplastic plasma cells resulted in pathological fractures (Figure 3).

The extensive BM replacement by neoplastic plasma cells, together with the skeletal destruction and the renal damage, resulted in anaemia in Sca1-MafB mice, as revealed by the significantly lower values of red blood cell count, haemoglobin, haematocrite, and mean corpuscular volume (Supplementary Figure 10). Therefore, these results show that Sca1-MafB mice exhibit reduced haematocrits, as it is typically found in MM patients.

Tumour plasma cells in Sca1-MafB mice are oligoclonal

In order to evaluate the clonal expansion of reprogrammed plasma cells, we evaluated the presence of an M-component (abnormal monoclonal immunoglobulin accumulation) in the serum of Sca1-MafB mice, as a result of the aberrant overproduction of immunoglobulin by the clonal neoplastic plasma cells. Western blottings of isoelectric focusing gels were performed (Supplementary Figure 11). Anti-IgG showed oligoclonal patterns in Sca1-MafB mice versus controls, as revealed by the comparison with a human hydridoma monoclonal control (Supplementary Figure 11A–C). The presence of this oligoclonal pattern becomes more evident in mice with a more advanced disease. The finding that the oligoclonality is only observed in the late stages of the disease is in agreement with the fact that Sca1 targets MafB expression to a significant percentage of the cells in the stem/progenitor compartment, implying that more than one cell-of-origin of the same type can contribute to the final reprogrammed plasma cells. This is further supported by the results from the analysis of V(D)J rearrangements in plasma cells from young Sca1-MafB mice (Supplementary Figure 12), which shows that there is not any predominant clone in this population. This indicates that malignant plasma cells arise from a population of equipotent and identically predisposed progenitors.

In humans, primary amyloidosis can be caused by a plasma cell neoplasm that secretes elevated titres of immunoglobulins to the serum, which afterwards deposit in various tissues. These amyloid deposits form a β-pleated sheet structure that binds Congo Red dye with a characteristic birefringence and induces changes in the fluorescence of the dye thioflavin T, and that is the direct responsible for the kidney failure frequently seen in the later stages of the disease (Jaffe *et al*, 2001). In this regard, the determination of the levels of the immunoglobulins in the serum of the Sca1-MafB mice showed that there is not a diffuse hypergammaglobulinemia although some of the transgenic mice showed elevated titres of specific immunoglobulins isotypes (Supplementary Figure 11D), as it is typically found in MM patients. In Sca1-MafB mice, paraprotein accumulation could be clearly detected in the spleen, liver, kidney, or intestine (Supplementary Figure 9A–C). The staining of the organs from diseased Sca1-MafB mice using thioflavin T or Congo Red proved the amyloid nature of the deposits, and provides further confirmation of the myelomatous disease affecting the animals (and responsible for the macroscopically visible organ alterations in the case of the kidneys).

Tumour plasma cells in Sca1-MafB mice display gene signatures analogous to those of human malignant plasma cells

The above results surprisingly show that it is possible to reprogram HS/PCs to malignant plasma cells by expressing the *MafB* oncogene in a restricted manner, exclusively in the $Sca1^+$ HS/PCs. The *MafB* oncogene, according to the mechanistic experimental design, is not expressed in more differentiated, $Sca1^-$, cell types, or even in the tumoral plasma cells themselves. This indicates that, somehow, programming of the malignant phenotype has already taken place at the stem cell level. This is, in fact, very similar to what it happens in the differentiation of normal haematopoietic lineages, where molecules like IL7-R or EpoR are only necessary at early developmental stages in order to program a given differentiation fate, but are not required afterwards, once the program has been established. In order to identify the molecular features of the pathological programming imposed by MafB in the stem cells in Sca1-MafB mice, we proceeded initially to compare the gene expression profiles of plasma cells ($B220^{low}CD138^{hi}FSC^{hi}SSC^{hi}$) of Sca1-MafB mice versus those from WT mice (Figure 4) using cDNA microarray analysis. Supervised analysis showed that, although the global gene expression pattern is quite similar between both populations, differences of ± 2 expression levels could be found (with a false discovery rate (FDR) = 6.65%) in a small percentage of genes, representing 0.8% of all the ones analysed in the array (Figure 4A). Further transcriptional analysis revealed that overexpression of genes known to be associated with human malignant plasma cells, including *Mum1/IRF4*, *Vegfa*, *IL6*, *Muc1*, *Fgf3*, *CD44*, and *Bcl2* (Casey *et al*, 1986; Williams, 1991; Matsumura and Tarin, 1992; Iida *et al*, 1997; Silva *et al*, 2001; van de Donk *et al*, 2005; Wei *et al*, 2006), were also expressed in tumour plasma cells of Sca1-MafB mice (Figure 4B). These results prove that Sca1-MafB tumour plasma cells share a genetic profile with human malignant plasma cells.

Maf-B target genes are not expressed in tumour plasma cells from Sca1-MafB mice

In these Sca1-MafB mice, oncogenic reprogramming takes place within the stem cell/progenitor population and, according to the experimental design, the *MafB* oncogene is not expressed in the tumour plasma cells themselves. However, it could be that MafB target genes continue to be expressed in the absence of MafB. These genes could be targets of a 'hit and run' mode of action in which MafB turns genes on in stem cells but is not required for maintaining their expression

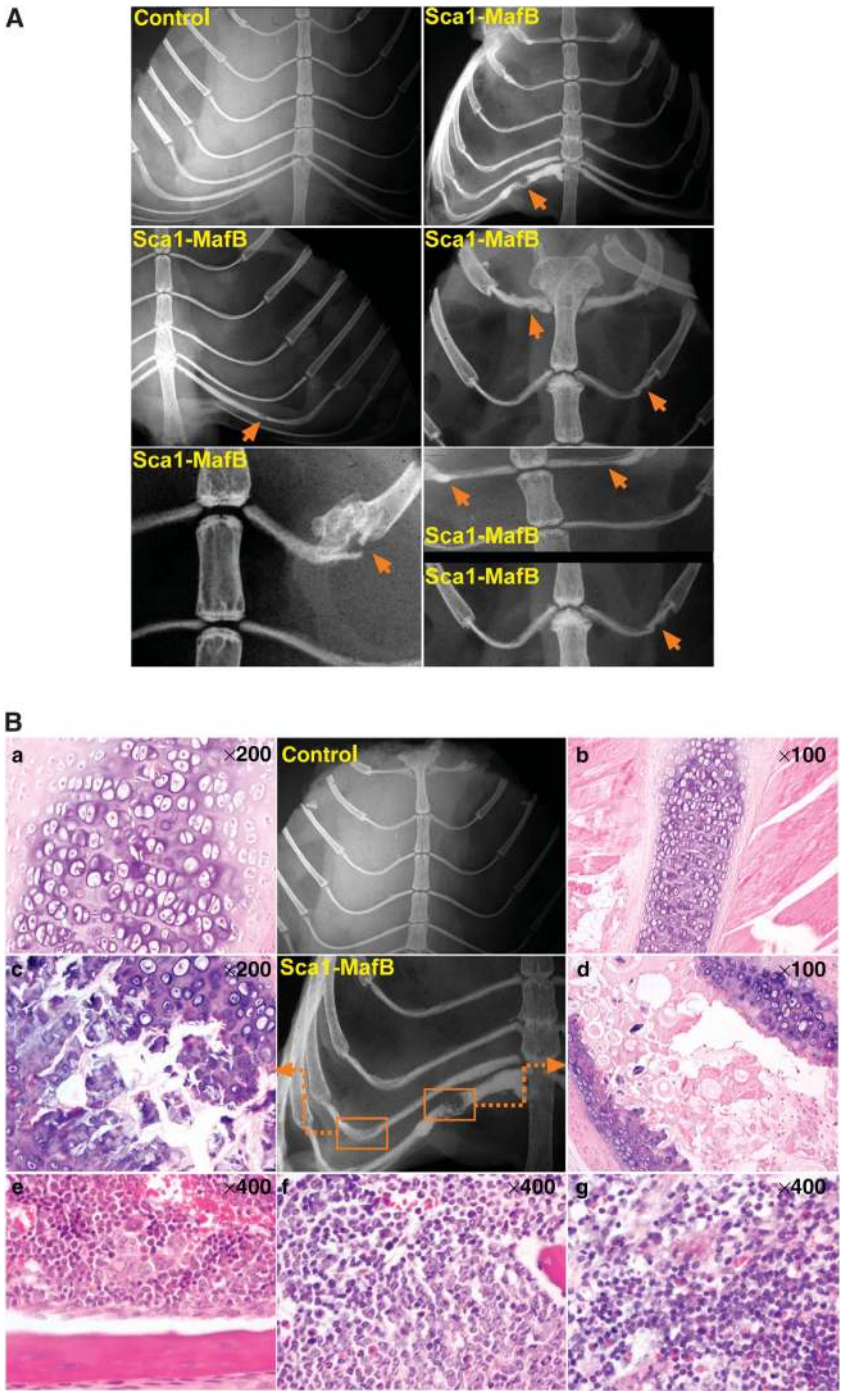
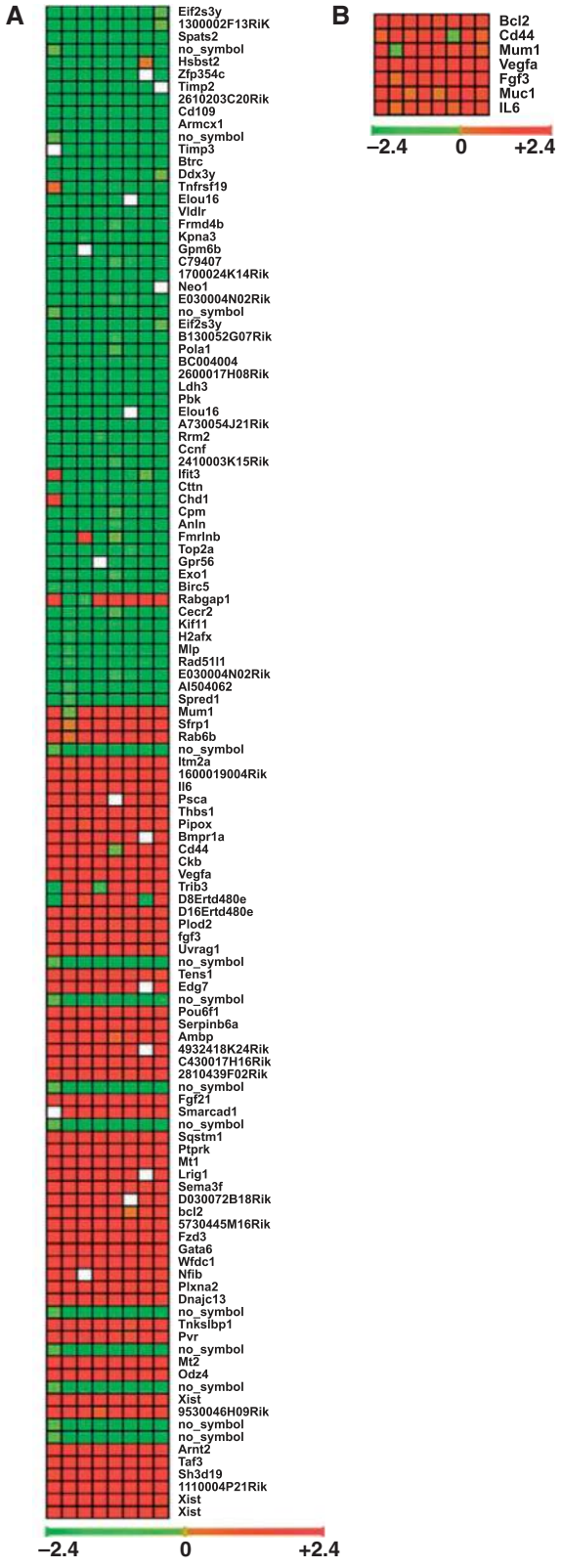


Figure 3 Bone lesions in Sca1-MafB mice. (**A**) Detection of osteolytic lesions and pathological fractures (indicated by orange arrows) by X-ray analysis of Sca1-MafB mice. A total number of nine mice (between 9 and 16 months old) were analysed, out of which five presented bone lesions. (**B**) Demonstration of specific plasma cell accumulation around the bone lytic areas in the Sca1-MafB mice after haematoxylin-eosin staining. Histologic appearance of ribs of both control (a, b) and Sca1-MafB mice (c, d). Plasma cells were not detected in the BM of ribs of control mice (e). However, plasma cells were present within the BM of ribs of Sca1-MafB mice (f). These plasma cells were infiltrating soft tissues around the ribs with osteolytic lesions (g). The image magnification is indicated.

at later stages of development. Remarkably, not only MafB expression is not detected in the plasma cells of Sca1-MafB mice, but what is more, neither MafB physiological target genes (*Ets-1*, *Hoxa3*, *Hoxb3*, *Gcg*, *Pax6*, *Maf*, *Fos*) (Kataoka *et al*, 1994; Sieweke *et al*, 1996; Kelly *et al*, 2000; Blanchi *et al*, 2003; Artner *et al*, 2007; Nishimura *et al*, 2008), nor the pathological downstream ones (*Ang*, *Blvra*, *Ccr1*, *Itgb7*, *Notch2*, *Cx3cr1*, *Ccnd2*, *Nuak1/Ark5*, *Ntrk2*, *Arid5a*, *Smarca1*, *Tlr4*, *Spp1*, *G6mb*, *Sfrp2*, *Tnfaip8*, *Dkk1*) (Zhan *et al*, 2002, 2003, 2006; van Stralen *et al*, 2009), were detected either (Figure 4; Supplementary Table II). It might appear then counter-intuitive and surprising that tumour plasma cells develop efficiently in these mice, since in actual human cancers all cancerous cells carry the oncogenic genetic

lesions, and not only the cancer stem cells. Nevertheless, tumour plasma cells arise in these mice, indicating that continuous expression of MafB is not critical for the generation of tumour plasma cells and suggesting a 'hands off' role for MafB in regulating tumour formation.



Neoplastic plasma cell commitment program is driven by the oncogene MafB

The results shown above prove that stem cell-restricted expression of MafB is enough to reprogram stem cells towards a neoplastic plasma cell fate. This is similar to what happens in other cases of normal fate programming or experimentally induced reprogramming, where the reprogramming factor(s) does not need to be present anymore once the initial fate-inducing change has taken place (e.g., induced pluripotency, see Discussion). Therefore, since the effects of MafB in the transgenic mice are restricted to the HS/PC compartment, we next sought for the differences in gene expression patterns between BM $Sca1^+Lin^-$ progenitors from Sca1-MafB mice when compared with those from WT mice. The $Sca1^+Lin^-$ population was isolated from 16 independent Sca1-MafB mice and was compared by gene expression profiling to normal $Sca1^+Lin^-$ cells from control mice. This provided an expression signature for the MafB-expressing HS/PCs, where a total of 4.3% of the genes (444 genes) analysed were upregulated or downregulated in Sca1-MafB $Sca1^+Lin^-$ cells (FDR $\leqslant$8.9%; Supplementary Figure 13; Supplementary Table III), revealing a common expression signature of MafB-expressing HS/PCs. The analysis identified a set of genes that are reproducibly differentially regulated in MafB-expressing HS/PCs versus controls, indicating that MafB expression does induce changes in the stem cell transcriptome that can affect cellular differentiation at later stages. Of particular interest was the upregulation of factors responsible for the terminal differentiation of B cells into plasma cells, like *Blimp1*, *Xbp1*, *Mum1*, and *Syndecan1/CD138* in MafB-expressing HS/PCs, results that were confirmed by real-time PCR analysis (Supplementary Figures 13 and 14). In order to rule out that these findings could be due to an enrichment of some specific subpopulation (e.g., common lymphoid progenitors, CLPs), we analysed and quantified both LSK and CLP compartments in transgenic and wild-type mice by flow cytometry (Supplementary Figure 4B and C). The results showed that the LSK population is increased in the transgenic mice. Using stainings including the Slam antibodies, we could identify that this change was due to an increase in the short-term HSC pool (Supplementary Figure 4B). On the contrary, the CLP population is decreased in transgenic mice (Supplementary Figure 4C). These

Figure 4 Molecular identity of plasma cells in Sca1-MafB mice. To molecularly characterize tumour plasma cells of Sca1-MafB mice, we compared the gene expression profiles of BM plasma cells ($B220^{low}CD138^{hi}FSC^{hi}SSC^{hi}$) purified from diseased Sca1-MafB mice versus those from the BM of control mice. Each gene (identified at right) is represented by a single row of coloured boxes; each mouse is represented by one single column. Data are displayed by a colour code. Red indicates overexpression in tumour plasma cells versus WT plasma cells, green indicates lower expression in the tumour versus WT plasma cells. (**A**) Tumour plasma cells are different from normal plasma cells. BM plasma cells were isolated from two different control mice and from eight different Sca1-MafB mice and independently extracted and amplified RNA from each aliquot for gene expression profiling. Genes for which signal intensities were up/down (threshold ± 2) regulated in Sca1-MafB plasma cells are shown. The ratios refer to the expression levels in Sca1-MafB plasma cells versus control plasma cells. (**B**) Gene expression pattern in plasma cells of Sca1-MafB mice ($n = 8$) of markers overexpressed in human malignant plasma cells. The expression levels of *Mum1/IRF4*, *Vegfa*, *IL6*, *Muc1*, *Fgf3*, *CD44*, and *Bcl2* in plasma cells of Sca1-MafB mice resemble those for plasma cells of human MM patients.

findings would therefore reinforce the interpretation of the results from the expression studies, and point towards the existence of a plasma cell-biased stem cell transformation signature that is dependent on the underlying MafB initiating mutation. These observations provide evidence indicating that different oncogenic proteins, when expressed in stem/progenitor cells can have selective impacts that depend on their intrinsic molecular properties. These results in turn provide a rationale for the strikingly consistent associations between different oncogenes and their corresponding cancer phenotypes (Sanchez-Garcia, 1997).

The existence of a MafB-induced stem cell signature suggests that deregulation of MafB can promote cancer through a specific oncogenic mechanism. Of particular interest was the fact that, as mentioned, neither the physiological MafB target genes (Kataoka *et al*, 1994; Sieweke *et al*, 1996; Kelly *et al*, 2000; Blanchi *et al*, 2003; Artner *et al*, 2007; Nishimura *et al*, 2008) nor the pathological downstream MafB targets (Zhan *et al*, 2002, 2003, 2006; van Stralen *et al*, 2009) were involved in the MafB-induced reprogramming of stem cells in Sca1-MafB mice (Supplementary Figure 13A). Therefore, the oncogenic activity of the MafB protein at the HS/PC level should most likely be due to alternative, 'non-canonical' functions arising in this specific cellular context (stem cells).

B cells in Sca1-MafB mice retain an epigenetic memory of their MafB-expressing HSCs of origin

Our results suggest that MafB imposes a gene regulatory state in stem cells that persists during haematopoiesis and mediates a latent tumour phenotype that only becomes overtly cancerous in the mature B-cell compartment, despite absence of MafB expression, suggesting a possible aberrant epigenetic programming. To confirm this hypothesis, and to determine the molecular mechanisms mediating MafB-induced reprogramming of stem cells into tumour plasma cells, we performed a complete, unbiased, and quantitative assessment of cytosine methylation from purified populations of primary cells. Using high-throughput reduced representation bisulphite sequencing (RRBS; Meissner *et al*, 2008; Bock *et al*, 2010; Gu *et al*, 2011), we generated DNA methylation maps for pooled FACS-isolated, $Sca1^+Lin^-$ fractions from *Sca1-MafB* mice and control mice, pooled $B220^+$ cells from *Sca1-MafB* mice and control mice, and pooled $Gr1^+Mac1^+$ cells from control mice (Figure 5). The heatmaps of the methylation values for the CpG islands and/or promoters allowed to cluster and distinguish the different cell types studied, indicating that the identified sites of increased/decreased methylation heterogeneity are indeed cell type-specific DNA-methylation regions (DMRs; Figure 5A and B). Several differentially hypermethylated and hypomethylated regions associated with promoters and with CpG islands were found in the comparisons and are shown in Figure 5C: control WT HS/PCs versus those from Sca1-MafB mice (MC1 comparison), mature B cells of Sca1-MafB mice versus WT control B cells (MC2 comparison), mature B cells of Sca1-MafB mice versus control WT HS/PCs (MC3 comparison), and WT control B cells versus WT control HS/PCs (MC4 comparison). The data from these comparisons are summarized in Table II and the full data are available in Supplementary Tables IV–XI. Remarkably, the majority of the methylation changes found in the MC1 comparison were also found in MC3 and MC4 experiments. However, in the MC2 comparison only 50% of this methylation pattern is maintained (Supplementary Tables XII–XVII; Supplementary Figure 15). Moreover, when plotting these identically methylated regions in a proportional Venn diagram, several overlapping regions appear (Figure 5D) showing 84 conserved methylation changes in regions associated with CpG islands and 52 conserved methylation changes in regions associated with promoters through all the comparisons. These observations support the notion that MafB imposes a specific epigenetic state in stem cells that persists in the mature B-cell compartment of Sca1-MafB mice, allowing us to identify an epigenetic DNA methylation memory of reprogrammed stem cells in the mature B cells of Sca1-MafB mice. Moreover, these results further indicate that differentiated tumour cells were derived from the Sca1-driven, *MafB*-expressing stem cells.

Discussion

MafB-induced reprogramming of the cellular identity of a normal stem/progenitor cell into a tumour plasma cell

Reprogramming is the rewiring of the transcriptional and epigenetic status of a cell to that of another cell type. It can be driven by different molecules, in physiological development, in tumoral development, or experimentally in the laboratory. Interestingly enough, the factors most commonly used in reprogramming to pluripotency have all been shown to play an oncogenic role in different contexts (Rowland and Bernards, 2005; Okita *et al*, 2007; Tanaka *et al*, 2007; Chen *et al*, 2008; Abollo-Jimenez *et al*, 2010; Castellanos *et al*, 2010; Sanchez-Garcia, 2010), thus further linking reprogramming to tumorigenesis. However, an essential question that remains to be answered in order to prove the reprogramming capacity of human oncogenes is whether normal progenitors can be reprogrammed to give rise to terminally differentiated tumour cells by the specifically associated oncogenes.

Our data provide proof-of-principle evidence of the fact that, at least in the context of MM development, HS/PCs can be directly reprogrammed into tumour plasma-like cells *in vivo* by an oncogene that in human patients is associated with plasma cell myeloma. This MafB-mediated reprogramming is, however, permissive in that it allows the normal differentiation of all haematopoietic cell types, and only reveals its malignant nature in the plasma cell compartment. In the oncogenic reprogramming model here presented, the reprogrammed $Sca1^+$ population can nevertheless complete a multistage differentiation pathway involving an initial commitment to the B-cell lineage and a subsequent differentiation to plasma cells. A similar scenario has been described in human CLL, where the propensity to generate malignant B cells has already been acquired at the HSC stage (Kikushige *et al*, 2011) and therefore, patient-derived HSCs showed an abnormal expression of lymphoid-related genes, presumably reflecting their cell-intrinsic pathologic priming into the lymphoid lineage. Notably, these HSCs did not show the typical chromosomal abnormalities of CLL (Kikushige *et al*, 2011), like we have shown for MM (Supplementary Figure 1).

But how does MafB instruct stem cells to give rise to a malignant plasma cell?

In spite of their well-proven involvement in some subsets of human MM, MAF proteins do not appear, on base of their

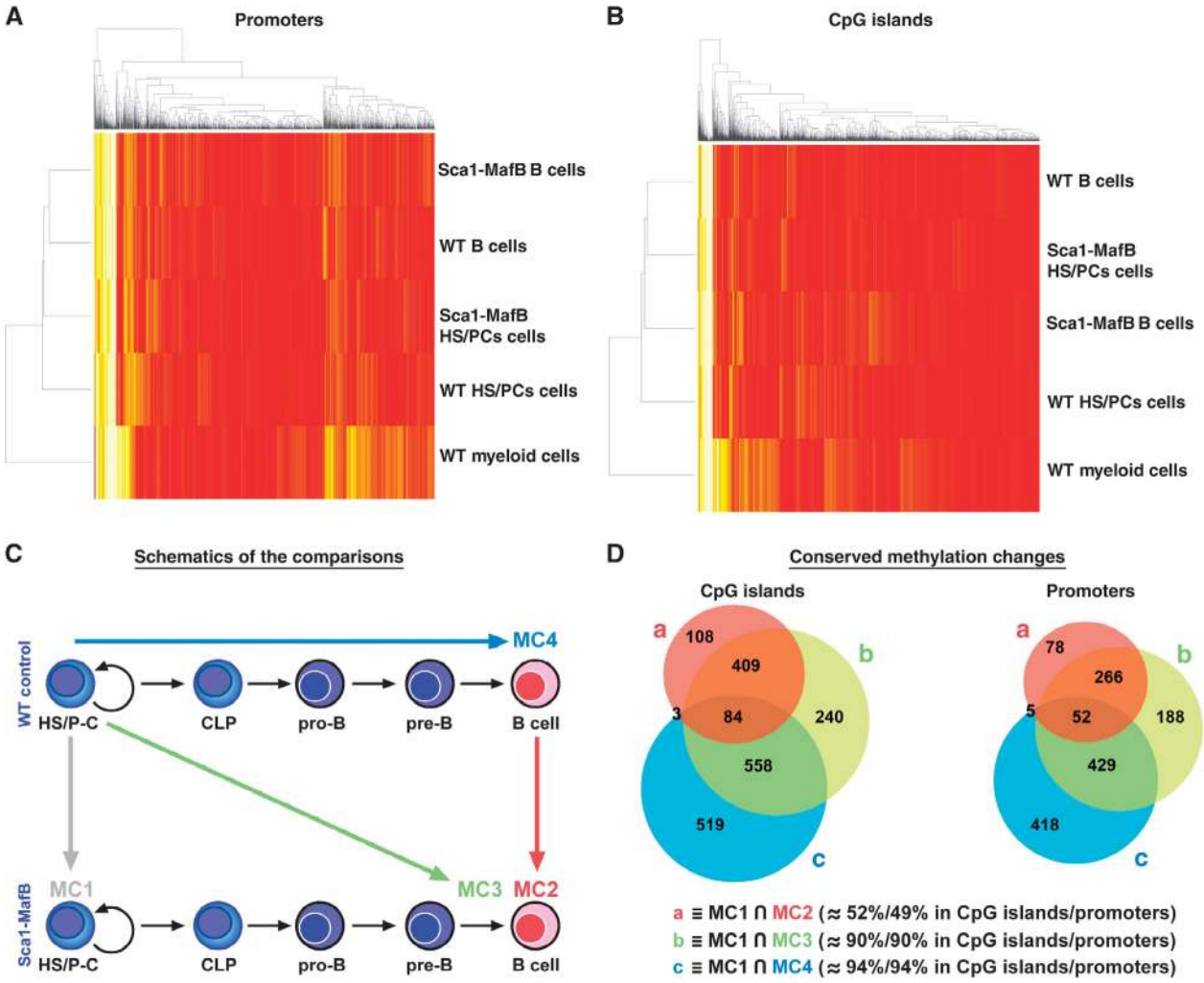


Figure 5 Genome-scale DNA methylation maps of stem and mature B cells of Sca1-MafB mice. (**A**, **B**) Heatmaps with hierarchical clustering. Samples were grouped based on differential methylation profiles for promoters (**A**) and CpG island loci (**B**). Yellow represents high levels of DNA methylation, and red represents low levels of DNA methylation. Methylation classification definitions: hypermethylated: 0–33% more methylated than reference; hypomethylated: 0–33% less methylated than reference; inconclusive: uneven or insufficient coverage for one or both samples; insignificant: statistically insignificant difference in methylation (*P*-value > 0.05); strongly hypermethylated: 33–100% more methylated than reference; strongly hypomethylated: 33–100% less methylated than reference. (**C**) Comparisons to detect methylation changes in HP/PCs of Sca1-MafB mice versus wild-type HS/PCs (MC1), mature B cells of Sca1-MafB mice versus wild-type mature B cells (MC2), mature B cells of Sca1-MafB mice versus control HS/PCs (MC3) and wild-type mature B cells versus wild-type HS/PCs (MC4). (**D**) Proportional Venn diagrams of conserved methylation changes in CpG islands and promoters associated regions. Each circle corresponds to the conserved changes between MC1 and MC2 (group a), MC1 and MC3 (group b) and MC1 and MC4 (group b). It also indicates the number of specific regions in each overlapping and non-overlapping zones of the diagrams.

Table II Number of differentially hypermethylated and hypomethylated regions associated to promoters or CpG islands in each comparison

	MC1	MC2	MC3	MC4
Regions associated with promoters				
Hypomethylated	858	2312	1927	1139
Hypermethylated	1390	1351	1674	1562
Regions associated with CpG islands				
Hypomethylated	1158	2958	2395	1363
Hypermethylated	1743	1484	1937	1948

physiological role in development, to be implicated in the major cellular functions associated to tumoral plasma cells. Indeed, MAF proteins are mostly inducers of terminal differentiation (Eychene *et al*, 2008), and MafB selectively restricts the sensitivity of stem cells to myeloid-inducing cytokines, limiting the development of myeloid lineages and contributing to the maintenance of a balanced lineage potential in the HSCs (Sarrazin *et al*, 2009). These activities cannot easily explain their oncogenic properties. Also, none of their physiological target genes has been found to be deregulated in Maf-transformed cells or to be involved in oncogenic processes (Eychene *et al*, 2008). In our experimental system, MafB targets are also not deregulated in MafB-reprogrammed stem cells. Therefore, the oncogenic activity of the MafB protein seems to be due to other, not yet described functions.

In order to identify the genes that are associated with MafB-induced reprogramming of stem cells, we performed a supervised analysis of the transcriptional profiles of HS/PCs

purified from Sca1-MafB mice and control mice. The data identified a set of genes that are reproducibly differentially regulated in MafB-targeted stem cells versus control stem cells, showing that Sca1-MafB-derived HSCs presented an abnormal expression of lymphoid-related genes, presumably reflecting their cell-intrinsic priming into the lymphoid lineage. Overall, these results show that (1) enforced MafB expression restricted to stem cells is all that is required to generate tumoral plasma cells in mice, therefore suggesting for the first time a role for stem/progenitor cells in the pathogenesis of plasma cell neoplasias; (2) HS/PCs transformed by MafB share a transformation-specific gene expression signature.

Among the genes present in the cancer stem cell signature there are some, like *Bmi1* or β-*catenin*, that are known to play relevant roles in the biology of both normal and cancer stem cells (Supplementary Figure 13B; Taipale and Beachy, 2001; Vonlanthen *et al*, 2001; Ema and Nakauchi, 2003; Lessard and Sauvageau, 2003) or to embryonic stem cells (Supplementary Figure 13C). Also, *Xbp1*, besides being essential for plasma cell development, has recently been found to significantly increase HSC activity in a functional screening based on retroviral transduction of $CD150^+CD48^-Lin^-$ mouse BM cells (Deneault *et al*, 2009). These findings provide a rationale of how reprogramming by MafB can induce changes from normal to pathological stem cells capable of giving rise to tumoral plasma cells.

In order to gain even more insight into the mechanism by which MafB-induced reprogramming of stem cells into tumour plasma cells, we generated *in-vivo* genome-scale maps of DNA methylation in both stem cells and mature B cells. We have found that a substantial number of CpG islands and promoters are specifically hypermethylated or hypomethylated in the stem cells of Sca1-MafB mice, setting a pattern inherited throughout B-cell development, similar to the DNA methylation changes taking place during somatic erythroid cell differentiation (Shearstone *et al*, 2011). Thus, these findings indicate that MafB oncogenic effect is driven in part by epigenetic mechanisms. In support of these findings, recent studies have shown that incorporation of targeted epigenetic agents to the standard chemotherapy is a promising approach to the treatment of relapsed paediatric ALL (Bhatla *et al*, 2012).

Until now, differentiated tumour cells have been obtained in mice mainly by targeting oncogene expression to mature cells. The results presented in this study demonstrate a novel molecular mechanism involved in tumour initiation, by showing that HS/PCs can be epigenetically reprogrammed to terminally differentiated tumour cells by defined oncogenes associated to human neoplasias (MafB, in this case, Figure 6). Our results also provide a proof-of-principle experimental model to understand how such a phenomenon could be possible also in human patients. We believe that our findings have critical implications for the understanding of the etiopathogenesis of MM and for the development of novel therapies for this disease.

Materials and methods

MM patient sample preparation and detection of chromosomal translocations by FISH

PC ($CD38^+CDE138^+$) and $CD34^+$ populations were sorted using a FACSAria flow cytometer (BDB) with CD38-PerCP-Cy5.5 (Pharmingen), CD138-FITC (Pharmingen) and CD34-APC (BDB) monoclonal antibodies, for further FISH analysis. Reanalysis of the sorted cells showed purity >95%. Interphase FISH studies for the detection of *IGH* rearrangements were carried out by means of LSI IGH dual-colour, break-apart rearrangement probe (Vysis, Downers Grove, IL, USA). Those MM samples with *IGH* translocations were explored for t(11;14)(q13;q32), t(4;14)(p16;q32), and t(14;16)(q32;q23) with the corresponding dual-colour, dual-fusion translocation probes from Abbott Molecular/Vysis, as previously described (Lopez-Corral *et al*, 2011).

Generation of Sca1-MafB transgenic mice

The Sca1-MafB vector was generated as follows: The 1.3-kb *Eco*RI-*Eco*RI fragment, containing the mouse *MafB* cDNA, was inserted into the *Cla*I site of the pLy6 vector (Miles *et al*, 1997) to generate the Sca1-MafB vector. The transgene fragment (Figure 1A) was excised from its vector by restriction digestion with *Not*I, purified for injection ($2\,ng\,ml^{-1}$) and injected into CBA × C57BL/6J fertilized eggs. Transgenic mice were identified by Southern analysis of tail snip DNA after *Eco*RI or *Hin*dIII digestion. *MafB* cDNA was used for detection of the transgene. A total of 44 transgenic animals and 20 control animals were used to define the phenotype. Two independent transgenic lines were generated (Figure 1B) and analysed and similar phenotypic features were seen in both.

EpiQuest library construction

EpiQuest libraries were prepared from 200 to 500 ng mouse genomic DNA obtained from primary cells ($Sca1^+Lin^-$ cells from BM, and $B220^+$ cells from PB) purified from *Sca1-MafB* ($n = 3$) mice and/or wild-type ($n = 10$) mice. The DNA was digested with 60 units of *Taq*I and 30 units of *Msp*I (NEB) sequentially. Size-selected *Taq*I-*Msp*I fragments (40–120 bp and 120–350 bp) were filled-in and 3′-terminal-A extended, extracted with Zymo Research DNA Clean and Concentrator(tm) kit. Ligation to pre-annealed adaptors containing 5′-methyl-cytosine instead of cytosine (Illumina) was performed using the Illumina DNA preparation kit and protocol. Purified, adaptor-ligated fragments were bisulphite treated using the EZ DNA Methylation-Direct(tm) Kit (Zymo Research). Preparative-scale PCR was performed and DNA Clean and Concentrator-purified PCR products were subjected to a final size selection on a 4% NuSieve 3:1 agarose gel. SYBR-green-stained gel slices containing adaptor-ligated fragments of 130–210 or 210–460 bp in size were excised. Library material was recovered from the gel (Zymoclean(tm) Gel DNA Recovery Kit) and sequenced on an Illumina GAIIx genome analyzer.

Sequence alignments and data analysis

Sequence reads from bisulphite-treated EpiQuest libraries were identified using standard Illumina base-calling software and then analysed using a Zymo Research proprietary computational pipeline.

Residual cytosines (Cs) in each read were first converted to thymines (Ts), with each such conversion noted for subsequent analysis. A reference sequence database was constructed from the 36-bp ends of each computationally predicted *Msp*I-*Taq*I fragment in the 40–220-bp size range. All Cs in each fragment end were then converted to Ts (only the C-poor strands are sequenced in the RRBS process). The converted reads were aligned to the converted reference by finding all 12-bp perfect matches and then extending to both ends of the treated read, not allowing gaps (reverse complement alignments were not considered). The number of mismatches in the induced alignment was then counted between the unconverted read and reference, ignoring cases in which a T in the unconverted read is matched to a C in the unconverted reference. For a given read, the best alignment was kept if the second-best alignment had two more mismatches, otherwise the read was discarded as non-unique. The methylation level of each sampled cytosine was estimated as the number of reads reporting a C, divided by the total number of reads reporting a C or T. A bioinformatics pipeline was used to score epigenetic alterations according to strength and significance, and links them to potentially affected genes. To that end, we collected a comprehensive set of regions of interest, which includes promoters, CpG islands, and repetitive elements. For each of these regions, the number of methylated and unmethylated CpG observations is determined, and a *P*-value is assigned using Fisher's exact test. Once all *P*-values are calculated, multiple-testing correction is performed separately

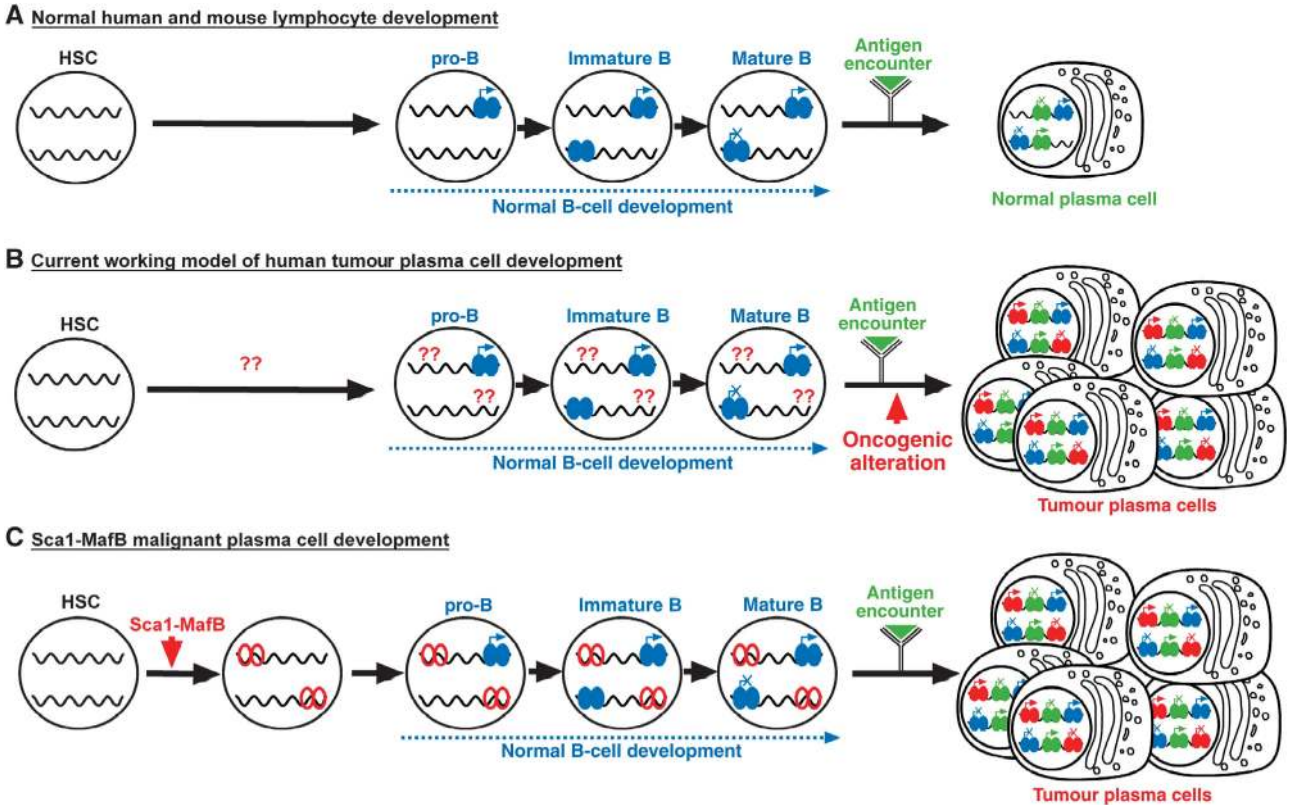


Figure 6 A model by which ectopic expression of MafB reprogrammes HS/PCs into tumour plasma cells. (**A**) Normal lymphoid development in human and mice. Blue circles represent normal gene regulatory events (activating or repressing) happening during B-lymphocyte development. Green circles represent normal gene regulation events happening during terminal differentiation to plasma cells, initially triggered by antigen recognition. (**B**) Current working model for the development of tumour plasma cells in humans. The existence of dormant alterations previous to the terminal differentiation is unknown. Currently, MafB effects (closed circles) are thought to occur in a plasma cell or in the final steps of the terminal differentiation from a B cell. The nature of the cancer cell-of-origin is therefore unknown. (**C**) Mechanism of tumour plasma cell development in Sca1-MafB transgenic mice. Open red circles represent latent epigenetic regulatory events caused by Sca1-driven expression of MafB. These epigenetic marks do not interfere with normal B-cell development, but become active (either activating or repressing) in the process of terminal differentiation, thus leading to the appearance of tumour plasma cell. According to this model, tumour plasma cell is the result of a cell reprogramming process (see text for details).

for each region type using the q-value method, which controls the FDR to be below a user-specified threshold (typically 10%). The software pipeline is implemented in Python (alignment processing module) and R (statistical analysis module).

Analysis of methylation levels

Once regions are classified, the information from all samples is processed in a different pipeline, also implemented in R. After filtering the inconclusive and insignificant values from each comparison, the common regions of control HS/PCs and those from Sca1-MafB mice (MC1) and (A) tumour B cells of Sca1-MafB mice versus B cells (MC2), (B) tumour B cells of Sca1-MafB mice versus control HS/PCs (MC3) and (C) B cells versus control HS/PCs (MC4) are extracted and combined (Supplementary Tables XII–XVII; Supplementary Figure 15) as shown in Figure 5C. This allows comparing the methylation levels for those common regions in each pair of comparisons.

With these combined data, the ratio of regions with conserved and changed methylation levels is calculated. Classes of conserved methylation levels and patterns of changed methylation levels are also determined. Overlapping and specific regions with conserved methylation levels in the three pairs of comparisons are represented in a Venn diagram. This analysis was carried out separately for promoters and for CpG islands.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (http://www.embojournal.org).

Acknowledgements

We thank all members of lab 13 at IBMCC for their helpful comments and constructive discussions on this project and Professor Jesús San Miguel for the samples and the critical reading of the manuscript. We are indebted to patients who gave samples. We are grateful to the Animal Care Facility staff members for their valuable help. We are grateful to Dr E Dzierzak for the Sca1 promoter and Dr Michael Sieweke for the mouse *Maf-B* cDNA. Research in ISG group was partially supported by FEDER and by MICINN (SAF2009-08803 to ISG), by Junta de Castilla y León (REF. CSI007A11-2 and Proyecto Biomedicina 2009-2010 to ISG), by MEC OncoBIO Consolider-Ingenio 2010 (Ref. CSD2007-0017), by NIH grant (R01 CA109335-04A1), by Sandra Ibarra Foundation, by Group of Excellence Grant (GR15) from Junta de Castilla y Leon, and the ARIMMORA project (FP7-ENV-2011, European Union Seventh Framework Program) and by proyecto en red de investigación en celulas madre tumorales en cancer de mama supported by Obra Social Kutxa y Conserjería de Sanidad de la Junta de Castilla y Leon. All Spanish funding is co-sponsored by the European Union FEDER program. CVD research is supported by Junta de Castilla y León (proyecto de investigación en biomedicina SAN/39/2010). ISG is an API lab of the EuroSyStem project. Research at CC's lab was partially supported by FEDER, Fondo de Investigaciones Sanitarias (PI080164), CSIC P.I.E., Junta de Castilla y León (SA060A09 and proyecto Biomedicina 2009-2010) and from an institutional grant from the Fundación Ramón Areces. IR-C and FA-J were supported by an FPU fellowship from the Spanish Ministerio de Ciencia e Innovacion. AO research is supported by a grant from the Instituto de Salud Carlos III, Ministerio de Sanidad y Consumo, Madrid, Spain (IISCIII-RTICC RD06/0020/0035-FEDER).

Author contributions: CV-D, CC, and ISG designed the study. CV-D, IG-H, EA-E, FA-J, RJ, AO, FJG-C, and MBG-C performed experiments. NG performed FISH analysis. NM and LMCG performed the isoelectric focusing. MCFC performed the X-ray studies. BP generated the transgenic mice. TF analysed mouse and human histopathology data. CV-D, EA-E, and TF performed IHC analyses in mouse samples. DA-L and JD analysed methylation data. XJ and IL determined Ig concentrations in serum and analysed data. CV-D, CC, and ISG coordinated and supervised the project. CV-D, CC, and ISG prepared figures and wrote the manuscript. All authors read and agreed the paper content.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Abollo-Jimenez F, Jimenez R, Cobaleda C (2010) Physiological cellular reprogramming and cancer. *Semin Cancer Biol* **20:** 98–106

Artner I, Blanchi B, Raum JC, Guo M, Kaneko T, Cordes S, Sieweke M, Stein R (2007) MafB is required for islet beta cell maturation. *Proc Natl Acad Sci USA* **104:** 3853–3858

Bhatla T, Wang J, Morrison DJ, Raetz EA, Burke MJ, Brown P, Carroll WL (2012) Epigenetic reprogramming reverses the relapse-specific gene expression signature and restores chemosensitivity in childhood B-lymphoblastic leukemia. *Blood* **119:** 5201–5210

Blair A, Sutherland HJ (2000) Primitive acute myeloid leukemia cells with long-term proliferative ability in vitro and in vivo lack surface expression of c-kit (CD117). *Exp Hematol* **28:** 660–671

Blanchi B, Kelly LM, Viemari JC, Lafon I, Burnet H, Bevengut M, Tillmanns S, Daniel L, Graf T, Hilaire G, Sieweke MH (2003) MafB deficiency causes defective respiratory rhythmogenesis and fatal central apnea at birth. *Nat Neurosci* **6:** 1091–1100

Bock C, Tomazou EM, Brinkman AB, Muller F, Simmer F, Gu H, Jager N, Gnirke A, Stunnenberg HG, Meissner A (2010) Quantitative comparison of genome-wide DNA methylation mapping technologies. *Nat Biotechnol* **28:** 1106–1114

Casey G, Smith R, McGillivray D, Peters G, Dickson C (1986) Characterization and chromosome assignment of the human homolog of int-2, a potential proto-oncogene. *Mol Cell Biol* **6:** 502–510

Castellanos A, Vicente-Duenas C, Campos-Sanchez E, Cruz JJ, Garcia-Criado FJ, Garcia-Cenador MB, Lazo PA, Perez-Losada J, Sanchez-Garcia I (2010) Cancer as a reprogramming-like disease: implications in tumor development and treatment. *Semin Cancer Biol* **20:** 93–97

Cobaleda C, Sanchez-Garcia I (2009) B-cell acute lymphoblastic leukaemia: towards understanding its cellular origin. *BioEssays* **31:** 600–609

Chen Y, Shi L, Zhang L, Li R, Liang J, Yu W, Sun L, Yang X, Wang Y, Zhang Y, Shang Y (2008) The molecular mechanism governing the oncogenic potential of SOX2 in breast cancer. *J Biol Chem* **283:** 17969–17978

Deneault E, Cellot S, Faubert A, Laverdure JP, Frechette M, Chagraoui J, Mayotte N, Sauvageau M, Ting SB, Sauvageau G (2009) A functional screen to identify novel effectors of hematopoietic stem cell activity. *Cell* **137:** 369–379

Ema H, Nakauchi H (2003) Self-renewal and lineage restriction of hematopoietic stem cells. *Curr Opin Genet Dev* **13:** 508–512

Eychene A, Rocques N, Pouponnot C (2008) A new MAFia in cancer. *Nat Rev Cancer* **8:** 683–693

Gemelli C, Montanari M, Tenedini E, Zanocco Marani T, Vignudelli T, Siena M, Zini R, Salati S, Tagliafico E, Manfredini R, Grande A, Ferrari S (2006) Virally mediated MafB transduction induces the monocyte commitment of human CD34+ hematopoietic stem/progenitor cells. *Cell Death Differ* **13:** 1686–1696

Gu H, Smith ZD, Bock C, Boyle P, Gnirke A, Meissner A (2011) Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nat Protoc* **6:** 468–481

Hamburger A, Salmon SE (1977) Primary bioassay of human myeloma stem cells. *J Clin Invest* **60:** 846–854

Hideshima T, Mitsiades C, Tonon G, Richardson PG, Anderson KC (2007) Understanding multiple myeloma pathogenesis in the bone marrow to identify new therapeutic targets. *Nat Rev Cancer* **7:** 585–598

Hurt EM, Wiestner A, Rosenwald A, Shaffer AL, Campo E, Grogan T, Bergsagel PL, Kuehl WM, Staudt LM (2004) Overexpression of c-maf is a frequent oncogenic event in multiple myeloma that promotes proliferation and pathological interactions with bone marrow stroma. *Cancer Cell* **5:** 191–199

Iida S, Rao PH, Butler M, Corradini P, Boccadoro M, Klein B, Chaganti RS, Dalla-Favera R (1997) Deregulation of MUM1/IRF4 by chromosomal translocation in multiple myeloma. *Nat Genet* **17:** 226–230

Jaffe ES, Harris NL, Stein H, Vardiman JW (2001) WHO Classification of Tumours. In *Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues*, Vol. 3, pp 142–154. Lyon, France: International Agency for Research on Cancer (IARC)

Kataoka K, Fujiwara KT, Noda M, Nishizawa M (1994) MafB, a new Maf family transcription activator that can associate with Maf and Fos but not with Jun. *Mol Cell Biol* **14:** 7581–7591

Kelly LM, Englmeier U, Lafon I, Sieweke MH, Graf T (2000) MafB is an inducer of monocytic differentiation. *EMBO J* **19:** 1987–1997

Kikushige Y, Ishikawa F, Miyamoto T, Shima T, Urata S, Yoshimoto G, Mori Y, Iino T, Yamauchi T, Eto T, Niiro H, Iwasaki H, Takenaka K, Akashi K (2011) Self-renewing hematopoietic stem cell is the primary target in pathogenesis of human chronic lymphocytic leukemia. *Cancer Cell* **20:** 246–259

Lessard J, Sauvageau G (2003) Bmi-1 determines the proliferative capacity of normal and leukaemic stem cells. *Nature* **423:** 255–260

Lim E, Vaillant F, Wu D, Forrest NC, Pal B, Hart AH, Asselin-Labat ML, Gyorki DE, Ward T, Partanen A, Feleppa F, Huschtscha LI, Thorne HJ, Fox SB, Yan M, French JD, Brown MA, Smyth GK, Visvader JE, Lindeman GJ (2009) Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat Med* **15:** 907–913

Lopez-Corral L, Gutierrez NC, Vidriales MB, Mateos MV, Rasillo A, Garcia-Sanz R, Paiva B, San Miguel JF (2011) The progression from MGUS to smoldering myeloma and eventually to multiple myeloma involves a clonal expansion of genetically abnormal plasma cells. *Clin Cancer Res* **17:** 1692–1700

Mak AB, Nixon AM, Moffat J (2012) The mixed lineage leukemia (MLL) fusion-associated gene AF4 promotes CD133 transcription. *Cancer Res* **72:** 1929–1934

Matsui W, Huff CA, Wang Q, Malehorn MT, Barber J, Tanhehco Y, Smith BD, Civin CI, Jones RJ (2004) Characterization of clonogenic multiple myeloma cells. *Blood* **103:** 2332–2336

Matsui W, Wang Q, Barber JP, Brennan S, Smith BD, Borrello I, McNiece I, Lin L, Ambinder RF, Peacock C, Watkins DN, Huff CA, Jones RJ (2008) Clonogenic multiple myeloma progenitors, stem cell properties, and drug resistance. *Cancer Res* **68:** 190–197

Matsumura Y, Tarin D (1992) Significance of CD44 gene products for cancer diagnosis and disease evaluation. *Lancet* **340:** 1053–1058

Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB, Gnirke A, Jaenisch R, Lander ES (2008) Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454:** 766–770

Miles C, Sanchez MJ, Sinclair A, Dzierzak E (1997) Expression of the Ly-6E.1 (Sca-1) transgene in adult hematopoietic stem cells and the developing mouse embryo. *Development* **124:** 537–547

Mitsiades CS, Mitsiades N, Munshi NC, Anderson KC (2004) Focus on multiple myeloma. *Cancer Cell* **6:** 439–444

Molyneux G, Geyer FC, Magnay FA, McCarthy A, Kendrick H, Natrajan R, Mackay A, Grigoriadis A, Tutt A, Ashworth A, Reis-Filho JS, Smalley MJ (2010) BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. *Cell Stem Cell* **7:** 403–417

Morito N, Yoh K, Fujioka Y, Nakano T, Shimohata H, Hashimoto Y, Yamada A, Maeda A, Matsuno F, Hata H, Suzuki A, Imagawa S,

Mitsuya H, Esumi H, Koyama A, Yamamoto M, Mori N, Takahashi S (2006) Overexpression of c-Maf contributes to T-cell lymphoma in both mice and human. *Cancer Res* **66:** 812–819

Morito N, Yoh K, Maeda A, Nakano T, Fujita A, Kusakabe M, Hamada M, Kudo T, Yamagata K, Takahashi S (2011) A novel transgenic mouse model of the human multiple myeloma chromosomal translocation t(14;16)(q32;q23). *Cancer Res* **71:** 339–348

Neering SJ, Bushnell T, Sozer S, Ashton J, Rossi RM, Wang PY, Bell DR, Heinrich D, Bottaro A, Jordan CT (2007) Leukemia stem cells in a genetically defined murine model of blast-crisis CML. *Blood* **110:** 2578–2585

Ng AP, Wei A, Bhurani D, Chapple P, Feleppa F, Juneja S (2006) The sensitivity of CD138 immunostaining of bone marrow trephine specimens for quantifying marrow involvement in MGUS and myeloma, including samples with a low percentage of plasma cells. *Haematologica* **91:** 972–975

Nishimura W, Rowan S, Salameh T, Maas RL, Bonner-Weir S, Sell SM, Sharma A (2008) Preferential reduction of beta cells derived from Pax6-MafB pathway in MafB deficient mice. *Dev Biol* **314:** 443–456

Nishizawa M, Kataoka K, Vogt PK (2003) MafA has strong cell transforming ability but is a weak transactivator. *Oncogene* **22:** 7882–7890

Okita K, Ichisaka T, Yamanaka S (2007) Generation of germline-competent induced pluripotent stem cells. *Nature* **448:** 313–317

Orfao A, Garcia-Sanz R, Lopez-Berges MC, Belen Vidriales M, Gonzalez M, Caballero MD, San Miguel JF (1994) A new method for the analysis of plasma cell DNA content in multiple myeloma samples using a CD38/propidium iodide double staining technique. *Cytometry* **17:** 332–339

Perez-Caro M, Cobaleda C, Gonzalez-Herrero I, Vicente-Duenas C, Bermejo-Rodriguez C, Sanchez-Beato M, Orfao A, Pintado B, Flores T, Sanchez-Martin M, Jimenez R, Piris MA, Sanchez-Garcia I (2009) Cancer induction by restriction of oncogene expression to the stem cell compartment. *Embo J* **28:** 8–20

Pouponnot C, Sii-Felice K, Hmitou I, Rocques N, Lecoin L, Druillennec S, Felder-Schmittbuhl MP, Eychene A (2006) Cell context reveals a dual role for Maf in oncogenesis. *Oncogene* **25:** 1299–1310

Rowland BD, Bernards R, Peeper DS (2005) The KLF4 tumour suppressor is a transcriptional repressor of p53 that acts as a context-dependent oncogene. *Nat Cell Biol* **7:** 1074–1082

Sanchez-Garcia I (1997) Consequences of chromosomal abnormalities in tumor development. *Annu Rev Genet* **31:** 429–453

Sanchez-Garcia I (2010) Getting to the stem of cancer. *Semin Cancer Biol* **20:** 63–64

Sarrazin S, Mossadegh-Keller N, Fukao T, Aziz A, Mourcin F, Vanhille L, Kelly Modis L, Kastner P, Chan S, Duprez E, Otto C, Sieweke MH (2009) MafB restricts M-CSF-dependent myeloid commitment divisions of hematopoietic stem cells. *Cell* **138:** 300–313

Shearstone JR, Pop R, Bock C, Boyle P, Meissner A, Socolovsky M (2011) Global DNA demethylation during mouse erythropoiesis in vivo. *Science* **334:** 799–802

Sieweke MH, Tekotte H, Frampton J, Graf T (1996) MafB is an interaction partner and repressor of Ets-1 that inhibits erythroid differentiation. *Cell* **85:** 49–60

Silva F, Carvalho F, Peixoto A, Seixas M, Almeida R, Carneiro F, Mesquita P, Figueiredo C, Nogueira C, Swallow DM, Amorim A, David L (2001) MUC1 gene polymorphism in the gastric carcinogenesis pathway. *Eur J Hum Genet* **9:** 548–552

Taipale J, Beachy PA (2001) The Hedgehog and Wnt signalling pathways in cancer. *Nature* **411:** 349–354

Tanaka Y, Era T, Nishikawa S, Kawamata S (2007) Forced expression of Nanog in hematopoietic stem cells results in a gammadeltaT-cell disorder. *Blood* **110:** 107–115

Tosi P, Gamberi B, Giuliani N (2006) Biology and treatment of multiple myeloma. *Biol Blood Marrow Transplant* **12:** 81–86

van de Donk NW, Lokhorst HM, Bloem AC (2005) Growth factors and antiapoptotic signaling pathways in multiple myeloma. *Leukemia* **19:** 2177–2185

van Stralen E, van de Wetering M, Agnelli L, Neri A, Clevers HC, Bast BJ (2009) Identification of primary MAFB target genes in multiple myeloma. *Exp Hematol* **37:** 78–86

Vicente-Duenas C, Barajas-Diego M, Romero-Camarero I, Gonzalez-Herrero I, Flores T, Sanchez-Garcia I (2012) Essential role for telomerase in chronic myeloid leukemia induced by BCR-ABL in mice. *Oncotarget* **3:** 261–266

Visvader JE (2011) Cells of origin in cancer. *Nature* **469:** 314–322

Vonlanthen S, Heighway J, Altermatt HJ, Gugger M, Kappeler A, Borner MM, van Lohuizen M, Betticher DC (2001) The bmi-1 oncoprotein is differentially expressed in non-small cell lung cancer and correlates with INK4A-ARF locus expression. *Br J Cancer* **84:** 1372–1376

Wei X, Xu H, Kufe D (2006) MUC1 oncoprotein stabilizes and activates estrogen receptor alpha. *Mol Cell* **21:** 295–305

Williams GT (1991) Programmed cell death: apoptosis and oncogenesis. *Cell* **65:** 1097–1098

Zhan F, Hardin J, Kordsmeier B, Bumm K, Zheng M, Tian E, Sanderson R, Yang Y, Wilson C, Zangari M, Anaissie E, Morris C, Muwalla F, van Rhee F, Fassas A, Crowley J, Tricot G, Barlogie B, Shaughnessy Jr J (2002) Global gene expression profiling of multiple myeloma, monoclonal gammopathy of undetermined significance, and normal bone marrow plasma cells. *Blood* **99:** 1745–1757

Zhan F, Huang Y, Colla S, Stewart JP, Hanamura I, Gupta S, Epstein J, Yaccoby S, Sawyer J, Burington B, Anaissie E, Hollmig K, Pineda-Roman M, Tricot G, van Rhee F, Walker R, Zangari M, Crowley J, Barlogie B, Shaughnessy Jr JD (2006) The molecular classification of multiple myeloma. *Blood* **108:** 2020–2028

Zhan F, Tian E, Bumm K, Smith R, Barlogie B, Shaughnessy Jr J (2003) Gene expression profiling of human plasma cell differentiation and classification of multiple myeloma based on similarities to distinct stages of late-stage B-cell development. *Blood* **101:** 1128–1140