cis-Regulatory Remodeling of the *SCL* Locus during Vertebrate Evolution[▽]

Berthold Göttgens,^{1†} Rita Ferreira,^{1†} Maria-José Sanchez,² Shoko Ishibashi,³ Juan Li,¹ Dominik Spensberger,¹ Pascal Lefevre,⁴ Katrin Ottersbach,¹ Michael Chapman,¹ Sarah Kinston,¹ Kathy Knezevic,¹ Maarten Hoogenkamp,⁴ George A. Follows,¹ Constanze Bonifer,⁴ Enrique Amaya,³ and Anthony R. Green^{1*}

Department of Haematology, Cambridge Institute for Medical Research, University of Cambridge, Cambridge CB2 0XY, United Kingdom¹; Centro Andaluz de Biología del Desarrollo, CSIC-Universidad Pablo de Olavide, Seville, Spain²; Healing Foundation Centre, Faculty of Life Sciences, University of Manchester, Manchester M13 9PT, United Kingdom³; and Section of Experimental Haematology, University of Leeds, St James's University Hospital, Leeds LS9 7TF, United Kingdom⁴

Received 27 July 2010/Returned for modification 21 August 2010/Accepted 3 October 2010

Development progresses through a sequence of cellular identities which are determined by the activities of networks of transcription factor genes. Alterations in *cis*-regulatory elements of these genes play a major role in evolutionary change, but little is known about the mechanisms responsible for maintaining conserved patterns of gene expression. We have studied the evolution of *cis*-regulatory mechanisms controlling the *SCL* gene, which encodes a key transcriptional regulator of blood, vasculature, and brain development and exhibits conserved function and pattern of expression throughout vertebrate evolution. *SCL cis*-regulatory elements are conserved between frog and chicken but accrued alterations at an accelerated rate between 310 and 200 million years ago, with subsequent fixation of a new *cis*-regulatory pattern at the beginning of the mammalian radiation. As a consequence, orthologous elements shared by mammals and lower vertebrates exhibit functional differences and binding site turnover between widely separated *cis*-regulatory modules. However, the net effect of these alterations is constancy of overall regulatory inputs and of expression pattern. Our data demonstrate remarkable *cis*-regulatory remodelling across the *SCL* locus and indicate that stable patterns of expression can mask extensive regulatory change. These insights illuminate our understanding of vertebrate evolution.

Spatiotemporal changes in gene expression patterns represent one of the main driving forces of evolution (reviewed in references 10 and 41), with morphological divergence being linked to mutations of specific *cis*-regulatory elements in a growing number of examples (3, 21, 29, 35, 42). Conversely, evolutionarily conserved features are thought to reflect conserved gene expression patterns, but little is known about the mechanisms responsible for such stability. Most evidence comes from invertebrate studies and points to disparate mechanisms. For example, in the urochordate *Ciona*, orthologous muscle-specific enhancers are highly conserved across 270 million years (7), whereas orthologues of the *Drosophila even-skipped* stripe 2 enhancer exhibit multiple differences in transcription factor binding sites and yet have maintained the same spatiotemporal activity (33, 34). In vertebrates, the *cis*-regulatory elements of many regulatory genes are not conserved in comparisons between mammals and lower vertebrates (1, 14, 49), and little is known about the mechanisms responsible for maintaining stable gene expression patterns.

The *SCL* gene encodes a basic helix-loop-helix (bHLH)

transcription factor with pleiotropic and nonredundant roles in the development of the hematopoietic (40, 43), vascular (20, 36), and central nervous systems (5, 20). Systematic dissection of *cis*-regulatory mechanisms operating at the murine *SCL* locus has identified multiple elements, each of which directs expression to a subdomain of the normal *SCL* expression pattern in transgenic mice. The two promoters (1a and 1b) have no hematopoietic activity (46), but three distal elements (kb -4, +19, and +40 with respect to the start of exon 1a) that are active in blood and/or endothelium have been identified. The -4 element directs expression to endothelial cells (46), the +19 to hematopoietic stem cells/progenitors (44, 45), and the +40 to erythroid cells (15). Each of these elements is characterized by a particular set of transcription factor binding sites. Promoter 1a (here referred as promoter) contains two functionally important GATA sites (4), the -4 element contains two ETS sites (24), the +19 enhancer contains an ETS/ETS/GATA motif (22), and the +40 element contains two GATA/E-box motifs (39). Mutation or deletion of a single one of these motifs leads to a loss of function of the enhancer (4, 22, 24, 39, 46).

The pattern of *SCL* expression is highly conserved throughout vertebrate evolution (23, 26, 28, 36, 46, 51). In this work, we have used stringent functional assays to characterize the *cis*-regulatory mechanisms responsible for maintaining the stability of this expression pattern. Our results reveal the existence of extensive *cis*-regulatory remodelling, which affects

* Corresponding author. Mailing address: Department of Haematology, Cambridge Institute for Medical Research, University of Cambridge, Cambridge CB2 0XY, United Kingdom. Phone: 004401223336820. Fax: 00441223762670. E-mail: arg1000@cam.ac.uk.

† Equal contribution.

▽ Published ahead of print on 18 October 2010.

multiple elements across the *SCL* locus. This unexpected plasticity has major implications for our understanding of vertebrate evolution.

MATERIALS AND METHODS

Sequence analysis. The human, mouse, and chicken *SCL* sequences and annotation were downloaded from our web server (http://hscl.cimr.cam.ac.uk/genomic_datasets.html). Frog *SCL* sequences were identified by BLAT analysis of the *Xenopus tropicalis* draft genome sequence at <http://genome.ucsc.edu/>. Global sequence alignments were performed using the lagan (http://lagan.stanford.edu/lagan_web/index.shtml) and chaos/dialign (<http://dialign.gobics.de/chaos-dialign-submission>) programs. Local sequence alignments were performed using ClustalW (48) installed locally and displayed using Genedoc (<http://www.psc.edu/biomed/genedoc/>). Genomic sequences were scanned for occurrences of transcription factor binding site consensus sequences using our own software program, TFBSsearch (13).

Transgenic mouse analysis. Chicken transgenic constructs were prepared using standard recombinant DNA technology. c*SCL* P/lac contained chicken sequences from −0.9 to +1.6 in exon 3 linked to a LacZ reporter gene. The c*SCL*-4/SV/lac, cSV/lac/*SCL*+19, and cSV/lac/*SCL*+40 constructs contained sequences from kb −2.1 to −0.9, +6.1 to +10.9, and +21.6 to +22.3, respectively, linked to the simian virus 40 (SV40) minimal promoter from pGL2-Promoter (Promega) and a LacZ reporter gene. Mouse transgenic constructs have been described (22, 39, 46). Transgenic mice were generated and analyzed as described previously (44). Both F0 embryos and stable transgenic mouse lines were used for the analysis. All mice were kept under specific-pathogen-free conditions, and all procedures were performed under the project license approved by the UK Home Office.

Transfection assays. Chicken core promoter fragments (188 and 142 bp) were amplified by PCR and cloned into the BglIII-XhoI sites of the pGL2-Basic luciferase vector (Promega). The ETS mutant was generated by PCR using specific primers containing the mutation (TTCC to CTCT for the first ETS site and GGAA to GGCC for the second ETS site). The chicken *SCL* +19 region was subcloned from SV/lac/*SCL*+19 into the SalI-BamHI sites of the pGL2-Promoter luciferase vector (Promega). The opossum 200-bp core promoter fragment and the platypus 250-bp core promoter fragment were PCR amplified from genomic DNA of opossum kidney cells and platypus BAC clone 200B24, respectively. Mutations were generated using the QuikChangeII XL site-directed mutagenesis kit (Stratagene). For luciferase assays, 416B cells were transfected and assayed as described previously (15).

In vivo DMS footprinting. The chicken cell lines DT40 (B cells), MEP (multipotent progenitors), and HD37 (erythroid) were cultured as described previously (8, 30), and *in vivo* dimethylsulfate (DMS) footprinting was performed as detailed previously (19). Briefly, living cells or genomic DNA was incubated at room temperature in 0.2% DMS solution in phosphate-buffered saline (PBS) for 5 min. The reaction was stopped with multiple washes with ice-cold PBS. Following cell lysis and DNA extraction, the DNA was cleaved with 0.1 M piperidine at 90°C for 10 min. Cleavage sites were analyzed by ligation-mediated PCR (LM-PCR). PCR products were labeled by primer extension using ³²P-labeled nested primers and were analyzed on 6% denaturing polyacrylamide gels as described previously (19). Linker and primer sequences are available on request. Experiments were performed on 3 separate occasions with material from 2 independent DMS-treated cell and genomic DNA preparations.

Chromatin immunoprecipitation. Chromatin immunoprecipitation (ChIP) was performed as described previously (15), using rabbit anti-*SCL* antibody (17) (a kind gift from Chris Drake, Charleston, SC) and nonspecific rabbit IgG (Sigma). Enrichment was determined using SYBR green real-time quantitative PCR (Stratagene) as described previously (15) and normalized against nonspecific IgG. Primers specific for the 3' untranslated region (3'UTR) of the chicken *SCL* gene and for the promoter of the hepatocyte-specific gene ApoVLDL2 (32) were used as negative controls.

Xenopus transgenics. The SV/Venus and SV/Venus/c*SCL*+19 constructs were generated by replacement of the LacZ reporter gene in SV/Lac and SV/Lac/c*SCL*+19 by the Venus sequence from Venus/pCS2 (38) (a kind gift from Atsushi Miyawaki, Saitama, Japan). Double transgenic *Xenopus laevis* embryos were generated using restriction enzyme-mediated integration (27, 31), using either SV/Venus or SV/Venus/c*SCL*+19 and γ -crystallin-mCherry, which was derived from γ -crystallin-green fluorescent protein (GFP) (6). The plasmids were linearized with SalI prior to the transgenesis protocol. Embryos expressing mCherry in the lens were fixed at stage 37/38, and Venus expression was assessed by whole-mount *in situ* hybridization using an antisense enhanced green fluores-

cent protein (EGFP) probe (47). Some of the embryos were allowed to develop to the larval stage (stage 47) in order to assess for Venus expression in the circulating blood.

RESULTS

Comparison of the *SCL* loci in different vertebrates. Comparative genome-based approaches are commonly used to identify *cis*-regulatory elements, based on the rationale that functionally important sequences remain unaltered through evolution while other sequences are allowed to change. Comparison of the human (*Homo sapiens*), murine (*Mus musculus*), chicken (*Gallus gallus*), frog (*Xenopus tropicalis*), and zebrafish (*Danio rerio*) *SCL* loci shows that flanking genes are largely conserved, with MAP17 situated immediately downstream of *SCL* in all four species (Fig. 1A). Long-range (global) DNA sequence alignment of the murine and human loci demonstrated that the promoter, as well as all three hematopoietic enhancers, was highly conserved. In contrast, alignment of the murine and chicken loci revealed significant sequence homology corresponding to the promoter but failed to identify peaks corresponding to the murine −4, +19, and +40 enhancers (Fig. 1B). Failure to identify homologues for regulatory elements by global sequence alignment is a well-known phenomenon in distal vertebrate comparisons (18). However, it is not clear whether this reflects the inability of sequence alignment algorithms to detect small but critical motifs in large stretches of otherwise highly diverged sequence or whether a large proportion of human/mouse regulatory elements are mammalian inventions with no lower vertebrate counterparts. Since the *SCL cis*-regulatory elements have been extensively characterized, we proceed to make use of the *SCL* locus as a model system to address this question.

The chicken *SCL* promoter has hematopoietic activity. Since the *SCL* promoter was the only *cis*-regulatory element to be identified by long-range DNA alignment, we initially focused on the structure and function of this conserved region. Comparative sequence analysis (human/mouse/chicken/frog) revealed an unexpected pattern (Fig. 2A). In addition to blocks of homology conserved between all 4 species (black shading), the chicken sequence was more closely related to the frog sequence (turquoise shading) than to the mammalian sequences. This observation was surprising, since chickens and mammals separated 310 million years ago whereas chickens and frogs have evolved separately for more than 350 million years (see Fig. S1 at http://hscl.cimr.cam.ac.uk/Gottgens_et_al_MCB_Supplemental_Information.pdf). Moreover, the additional stretches of conservation restricted to chicken and frog contained ETS and GATA motifs reminiscent of the combination found in the mammalian +19 hematopoietic/endothelial enhancer (22). These data suggested that the promoter element of lower vertebrates might have additional functionality that compensates for the apparent lack of distal hematopoietic/endothelial enhancers.

We have shown previously that individual murine *SCL cis*-regulatory elements reproducibly direct *lacZ* expression to subdomains of the normal *SCL* expression pattern (see Fig. S2 at http://hscl.cimr.cam.ac.uk/Gottgens_et_al_MCB_Supplemental_Information.pdf) (15, 22, 24, 39, 46). Therefore, we used transgenic reporter assays to compare the functions of the mouse and chicken promoter regions. A murine promoter construct directed

FIG. 1. Structure and sequence conservation of vertebrate *SCL* locus. (A) Structures of the human (H), mouse (M), chicken (C), frog (F), and zebrafish (Z) *SCL* loci. All loci are drawn to the same scale (numbers represent kb). Boxes represent exons, and arrows indicate gene orientation. The gray shading indicates the region of synteny. (B) Vista conservation plots of a global DNA sequence alignment between the mouse and human (top) or mouse and chicken (bottom) *SCL* loci. Exons are depicted as boxes, conserved coding regions are in blue, and conserved noncoding regions are in pink. Red arrowheads indicate the locations of murine *SCL* regulatory elements.

expression to the midbrain and spinal cord (46), as did an equivalent chicken construct. However, the chicken construct was also active in endothelial cells, aortic clusters, and a subpopulation of hematopoietic cells in fetal liver (Fig. 2B and Table 1; see also Fig. S3 and S4A at http://hscl.cimr.cam.ac.uk/Gottgens_et_al_MCB_Supplemental_Information.pdf). Consistent with these results, the chicken core promoter was highly active in stable transfection assays using the murine hematopoietic progenitor cell line 416B, whereas the equivalent murine core promoter was minimally active (Fig. 2C). Deletion of a 5' fragment, containing the additional

ETS and GATA motifs conserved between chicken and frog, did not affect activity in 416B cells, but mutation of the two 3' ETS sites abolished activity. Consistent with the transfection data, *in vivo* footprinting studies demonstrated that the 3' GATA/GATA/ETS/ETS motifs were indeed occupied in the *SCL*-expressing chicken hematopoietic cell lines MEP (hematopoietic progenitors) and HD37 (erythroblast) but not in the *SCL*-negative cell line DT40 (B cell) (Fig. 2D; see also supplemental Fig. 4B at the above URL). Moreover, the promoter of murine *LYL-1*, an *SCL* paralogue, contains GATA and ETS sites and is active in hema-

TABLE 1. Activity of mouse and chicken *SCL* regulatory elements in transgenic embryos

SCL <i>cis</i> -regulatory element	Species (reference)	Expression ^b					
		Midbrain	Spinal cord	Blood	Fetal liver	Dorsal aorta clusters	Endothelium
None ^c	NA ^a	0/21 (–)	2/21 (++) ^d	0/21 (–)	0/21 (–)	ND	1/21 (++) ^d
Promoter	Mouse (46)	7/8 (+++)	7/8 (+++)	0/8 (–)	0/8 (–)	0/8 (–)	0/8 (–)
	Chicken	6/7 (+++)	6/7 (+++)	4/7 (+)	4/5 (+)	3/5 (++)	3/5 (+)
–4	Mouse (24)	0/7 (–)	0/7 (–)	0/7 (–)	4/7 (++)	4/7 (+++)	4/7 (+++)
	Chicken	0/9 (–)	0/9 (–)	2/9 (+)	3/3 (++)	0/3 (–)	2/3 (+)
+40	Mouse (39)	9/10 (+++)	0/10 (–)	9/10 (+++)	9/10 (++)	2/4 (+)	0/4 (–)
	Chicken	11/11 (+++)	3/11 (++)	2/11 (++)	4/5 (++)	1/5 (++)	1/5 (++)
+19	Mouse (25)	0/7 (–)	0/7 (–)	0/7 (–)	5/7 (++)	2/2 (++) ^e	5/7 (+++)
	Chicken	0/6 (–)	0/6 (–)	5/5 (+++)	5/5 (++)	1/5 (+)	1/5 (+)
+19 GATA mutant	Chicken	0/8 (–)	0/8 (–)	0/8 (–)	0/4 (–)	0/4 (–)	0/4 (–)
+19 E-box mutant	Chicken	0/11 (–)	0/11 (–)	0/11 (–)	0/4 (–)	0/4 (–)	0/4 (–)

^a NA, not applicable.^b Numbers indicate embryos showing expression in the specific tissue/total number of transgenic embryos analyzed (scored by PCR). Symbols in parentheses signify no (–) or a low (+), medium (++), or strong (+++) level of LacZ expression. ND, not determined. Expression in brain, spinal cord, and blood was scored in whole-mount embryos, while expression in fetal liver, dorsal aorta clusters, and endothelium was scored in sagittal sections. The results for mouse constructs have been published previously (25, 24, 39, 46). Data in this table represent results for independent F0 embryos together with stable transgenic lines. Transgenic lines were generated by mating F0 transgenic males with (C57BL/6 × CBA) F1 females: cSCL P (2 lines; L6048 and L6058) or cSCL +19 (2 lines; L3699 and L4000). All stable transgenic lines gave expression patterns similar to the ones observed in F0 embryos.^c The SV40 minimal promoter was used to drive expression of the LacZ reporter gene.^d Ectopic LacZ expression was observed in a small minority of transgenic embryos carrying the SV40 minimal promoter transgene.^e Only two of the seven transgenic embryos were sectioned, allowing confirmation of expression in the dorsal aorta clusters; however, other mSCL +19 constructs show LacZ expression in dorsal aorta clusters (22).

functional motifs (13, 16). In transgenic reporter assays, the murine –4 enhancer was strongly active in all endothelial cells and in dorsal aorta clusters and targeted a proportion of fetal liver hematopoietic cells. In contrast, although activity of the chicken element was similar in fetal liver, endothelial activity was detectable only in rare cells (Fig. 3B and Table 1; see also Fig. S3 at http://hscl.cimr.cam.ac.uk/Gottgens_et_al_MCB_Supplemental_Information.pdf).

A similar approach identified a pair of GATA/E-box composite motifs downstream of the chicken and frog *MAP17* genes and in the same orientation as those in the mouse +40 enhancer (Fig. 3C). The putative chicken element was functional in transgenic analyses and gave rise to midbrain staining in a pattern identical to that of the mouse element. The mouse +40 enhancer and its chicken orthologue exhibited similar activities in fetal liver, but the expression of the chicken element in circulating erythrocytes was much lower than that of mouse +40 (9/10 and 2/11, respectively) (Fig. 3D and Table 1; see also Fig. S3 at http://hscl.cimr.cam.ac.uk/Gottgens_et_al_MCB_Supplemental_Information.pdf). These data demonstrate that the chicken *SCL* locus contains orthologous enhancers functionally related to the murine –4 and +40 elements.

The murine *SCL* +19 enhancer and its chicken orthologue use different regulatory codes. Identifying an orthologue for the mouse +19 enhancer was even more challenging. This mouse element contains a composite ETS/ETS/GATA motif with each individual binding site required for enhancer function (22), yet none of the 8 GATA motifs in the chicken *SCL* 3' flanking sequence (4.8 kb from the chicken *SCL* poly(A) site to the equivalent of mouse *SCL* +23 [23]) was accompanied by two

ETS motifs within any 50-nucleotide window (see Fig. S5A at http://hscl.cimr.cam.ac.uk/Gottgens_et_al_MCB_Supplemental_Information.pdf). However, one of these 8 GATA sites showed extended conservation (10/11 conserved nucleotides) with the mammalian +19 GATA site (Fig. 4A; see also Fig. S5B at the above URL). This was the only 10/11 match to the extended GATA motif identified within the syntenic region of the chicken locus, a 30-kb fragment ranging from the poly(A) site of the *SIL* gene to the first exon of the *CYP4X* gene. Furthermore, the region 3' of the frog *SCL* gene contained a GATA site with an 8/11 match to this particular chicken GATA motif, and strikingly, both chicken and frog GATA sites were preceded by an E-box motif in an 8-bp block of sequence identity (Fig. 4A). Transgenic analysis confirmed the presence of a functional chicken enhancer. Whereas the mouse +19 enhancer was active in fetal liver, endothelium, and dorsal aorta clusters, the chicken enhancer was active mainly in circulating erythroid cells and fetal liver (Fig. 4B and Table 1; see also Fig. S3 at http://hscl.cimr.cam.ac.uk/Gottgens_et_al_MCB_Supplemental_Information.pdf). The GATA and E-box motifs were both required for full activity of the chicken enhancer in stable transfection assays (Fig. 4C) and in transgenic embryos (Table 1 and Fig. 4D). *In vivo* footprinting and chromatin immunoprecipitation (Fig. 4E and F) demonstrated occupancy of the E box and binding of the *SCL* protein to the chicken element in chicken hematopoietic cells. Consistent with the higher expression in circulating erythroid cells, the occupancy of the E box by *SCL* was higher in HD37 (erythroid) cells than in MEP (progenitor) cells. The bipartite GATA/E-box motif is commonly found in erythroid cell-specific regulatory elements (50),

where it is bound by a multiprotein complex containing *SCL* and *GATA-1*. Due to the unavailability of antibody reagents against relevant chicken transcription factors other than chicken *SCL*, binding by chicken *GATA-1* to the bipartite motif could not be proven directly. However, the cumulative evidence presented here strongly suggests that the chicken +19 enhancer operates through the typical bipartite *GATA/E-box* motif bound by *SCL* and *GATA-1* in erythroid cells, because (i) the *GATA* site is conserved between chicken and frogs, (ii) the endogenous chicken *SCL* binds to the *E-box* portion of the bipartite motif, and (iii) mutation studies showed that both the *E-box* and *GATA* motifs were essential for erythroid activity of the chicken +19 enhancer when tested in both cell lines and transgenic mice.

The distinctive activity of the chicken +19 equivalent was confirmed in transgenic frogs. We have previously demonstrated that the mouse +19 enhancer directs expression in transgenic frogs to cells in the dorsolateral plate that are fated to give rise to hematopoietic progenitors and endothelium (22). In marked contrast, the chicken orthologue of the +19 element was active in the ventral blood islands of stage 37/38 embryos, the site of primitive erythropoiesis, and in the vast majority of circulating erythroid cells in stage 47 tadpoles (Fig. 5; see also Movies S1 and S2 at <http://hscl.cimr.cam.ac.uk/Video%20S1.mov> and <http://hscl.cimr.cam.ac.uk/Video%20S2.mov>, respectively). The results obtained using both transgenic mice and transgenic frogs are concordant with the erythroid activity of the chicken element in mice and also with the known erythroid activity of the *E-box/GATA* motif. Our data therefore demonstrate that the chicken *SCL* locus does contain a 3' enhancer orthologous to the mouse +19 element but, remarkably, the two enhancers use different regulatory codes, the *GATA/E-box* erythroid code (50) and the hematopoietic stem cell/endothelial *ETS/ETS/GATA* code (25).

Timing of *cis*-regulatory remodelling: the birth of a hematopoietic stem cell enhancer. Our results demonstrate that the *cis*-regulatory elements regulating *SCL* expression in chicken are strikingly different from those operating in eutherian mammals. To better define the evolutionary timing of these changes, we cloned, sequenced, and annotated a BAC carrying the platypus (*Ornithorhynchus anatinus*) *SCL* locus (see Fig. S6 at http://hscl.cimr.cam.ac.uk/Gottgens_et_al_MCB_Supplemental_Information.pdf) and compared the *SCL* loci of platypus and opossum to those of eutherian mammals. The region between *SCL* and *MAP17* in both platypus and opossum (*Monodelphis domestica*) contained the conserved *ETS/ETS/GATA* motif found in the +19 enhancer homologues of eutherian mammals (Fig. 6A). In contrast, the *SCL* promoters of platypus and opossum contained two of the four *ETS* motifs found in chicken and frog

(Fig. 6B). Moreover, compared to the chicken and murine promoters, the platypus and opossum promoters showed intermediate activity in stable transfection assays in the 416B hematopoietic progenitor cell line, and mutation of the 5' binding sites (*ETS*, *ETS*, and *GATA*) of the platypus promoter reduced its activity to the level of the mouse element's (Fig. 6C). These data indicate that the emergence of the mammalian +19 structure, which directs expression to the vast majority of hematopoietic stem cells (44), preceded the complete loss of hematopoietic function in the *SCL* promoter and in a window between 310 million and 200 million years ago.

DISCUSSION

Over the last 2 decades, the concept of alteration in *cis*-regulatory elements, rather than coding sequences, has gained increasing support as a key driver for the evolution of morphological differences (reviewed in references 10 and 41). The majority of studies so far have focused on the analysis of *cis*-regulatory elements of invertebrate genes, alterations of which cause striking morphological changes in closely related species (3, 21, 29, 35, 42), with only a few examples in vertebrates (3). Compared to the evolution of altered gene expression patterns, much less is known about the mechanisms responsible for highly conserved patterns of expression. The latter might reflect highly conserved *cis*-regulatory mechanisms; some invertebrate data support this idea (3, 21, 29, 35, 42), but little is known about the mechanisms operating in vertebrates. In this study, we have used comparative genomic sequence analysis to identify the chicken *SCL* *cis*-regulatory elements. Where long-range DNA alignments failed to identify conserved regions, we used our previous characterization of transcription factor binding sites essential for the activity of murine *SCL* regulatory elements to design targeted searches. To analyze the function of the chicken *cis*-regulatory elements and avoid possible functional redundancy, each element was analyzed independently. The function of each element was analyzed *in vitro*, using luciferase assays, in different hematopoietic cell lines and *in vivo* using transgenic mice and frogs. Both approaches showed that despite the conservation of gene expression patterns between mice and chickens, each independent element shows considerable functional differences compared with its mouse orthologue.

Several aspects of our data are of general significance. First, they demonstrate that a highly conserved expression pattern may hide unexpected *cis*-regulatory plasticity (Fig. 6D). The functions of individual *SCL* elements are different in mammals and lower vertebrates. In some cases, the altered function reflects binding site turnover between distinct elements—the mammalian promoter has lost *ETS* sites, but these have been

FIG. 3. The murine *SCL* -4 and *SCL* +40 enhancers are functionally related to the orthologous chicken enhancers. (A) Four-way sequence alignment of the *SCL* -4 region between human, mouse, chicken, and frog (shading as for Fig. 2A). Putative transcription factor binding sites are indicated. (B) E11.5 to E12.5 F0 transgenic embryos generated with the *SCL* -4 region from mouse (left panel) or chicken (right panel) stained with X-Gal. For each embryo, sagittal sections through the fetal liver (FL) and dorsal aorta (DA) are shown. (C) Four-way sequence alignment of the *SCL* +40 region between human, mouse, chicken, and frog (shading as for Fig. 2A). Putative transcription factor binding sites are indicated. (D) E11.5 to E12.5 F0 transgenic embryos generated with the *SCL* +40 region from mouse (left panel) or chicken (right panel) stained with X-Gal. For each embryo, sagittal sections through the fetal liver (FL) and dorsal aorta (DA) are shown.

FIG. 4. The murine SCL +19 enhancer and its chicken orthologue use different regulatory codes. (A) Four-way sequence alignment of the SCL +19 region between human, mouse, chicken, and frog (shading as for Fig. 2A). Putative transcription factor binding sites are indicated. (B) E11.5 F0 transgenic embryos generated with the SCL +19 region from mouse (1.1 kb; left panel) or chicken (1.5 kb; right panel) stained with X-Gal. For each embryo, sagittal sections through the fetal liver (FL) and dorsal aorta (DA) are shown. (C) Luciferase assay of 416B cells stably transfected with chicken SCL +19 constructs. A schematic representation of the constructs is shown on the left, with luciferase assay results on the right. The activity of a 1.5-kb fragment containing the SCL +19 enhancer was compared with the same fragment with mutated E-box (CAGATG to GAGAGT), GATA (GATA to GTGC), or both. Luciferase data are presented as relative luciferase activity increases over that of the minimal SV40 promoter negative control. Data shown represent the means from at least 4 independent pools of transfected cells. (D) In contrast to the case with wild-type chicken SCL +19, transgenic embryos generated with chicken SCL +19, where the E-box (SV/lac/cSCL+19 E-box mutant) or GATA (SV/lac/cSCL+19 GATA mutant) sites were mutated, show no expression in fetal liver (FL), dorsal aorta (DA), or endothelium (H, heart). (E) DMS *in vivo* footprinting shows protection of the E box in the HD37 erythroid cell line but not in the nonerythroid cell lines DT40. Weak protection is also observed in the MEP cell line. Naked DNA from MEP (C-MEP) and HD37 (C-HD37) was used as a control. A representative experiment is presented. Protected sites are indicated by asterisks. (F) SCL is bound to the chicken +19 element in MEP and HD37 SCL-expressing chicken cell lines. The graph shows results of a ChIP assay using anti-chicken SCL antibody. Results are shown as fold increase over IgG. The ApoVLDL2 promoter (hepatocyte-specific gene) and SCL 3'UTR were used as negative controls.

acquired by the mammalian +19 enhancer. In acquiring its ETS sites, the mammalian +19 enhancer has lost an E box present in the equivalent ancestral enhancer and has switched from an erythroid element to a hematopoietic stem cell/endothelial enhancer. The ancestral SCL locus contained two enhancers with erythroid GATA/E-box motifs (orthologous to the murine +19 and +40 enhancers), and this redundancy may have permitted the +19 equivalent to lose its E box without deleterious consequences. In the mouse, the +40 is the major element driving activity to the circulating erythroid cells,

whereas in the chicken, the element corresponding to the +19 fulfils this role, raising the possibility that the chicken +40 orthologue functions primarily in the midbrain. Despite the multiple differences in the *cis* regulation of the mouse and chicken, the net effect is a constancy of regulatory inputs and spatiotemporal transcriptional output. However, this constancy conceals significant *cis*-regulatory alterations throughout the locus, with orthologous elements from different species exhibiting structural and functional differences. These observations also demonstrate that fixation of mammal-specific reg-

FIG. 5. The chicken orthologue to the murine *SCL* +19 enhancer drives reporter gene expression to the ventral blood island in *Xenopus* embryos. The chicken *SCL* +19 enhancer drives reporter gene expression to the ventral blood island in *Xenopus* transgenics. Double transgenic *Xenopus laevis* embryos were generated by the restriction enzyme-mediated integration method using SV/Venus or SV/Venus/c*SCL*+19 and γ -crystallin-mCherry. Embryos expressing mCherry in the lens were fixed at stage 37/38, and Venus expression was detected by whole-mount *in situ* hybridization using an antisense EGFP probe. Endogenous *SCL* mRNA was detected in wild-type embryos at the same stage, using an antisense *SCL* probe (right panel). The bottom panel shows transverse sections of the same embryos. Arrowhead, ventral blood island.

ulatory elements (2, 37) is not necessarily accompanied by the acquisition of new transcriptional inputs or changes in the overall expression pattern of a gene but may reflect “rewiring” within a locus. It is unclear how new transcription factor binding sites “appear” within an element, although it has recently been hypothesized that in addition to conventional evolutionary processes, specific binding sites may also be transferred between elements through recombination-based mechanisms (9).

Secondly, the data emphasize that when using comparative genomic sequence analysis to identify regulatory elements, an absence of evidence is not evidence of absence. Sequence alignments were unable to detect enhancers that could be identified given detailed knowledge of functionally important motifs. The +19 enhancer provides a particularly dramatic example, where a few conserved nucleotides flanking a consensus motif provided the clue to locating the chicken equivalent. Conversely, sequence conservation between orthologous regulatory elements may not equate to conserved function. The *SCL* promoter region exhibits significant sequence conservation throughout vertebrate evolution, and yet the mammalian promoter lacks hematopoietic/endothelial activity displayed by the *SCL* promoter from lower vertebrates.

Thirdly, our results have revealed an unexpected pattern of sequence conservation: the chicken sequence for each *SCL* regulatory element is more closely related to its frog orthologue than to its mammalian equivalent. This unexpected pattern of sequence conservation is not a general feature of the *SCL* locus, since alignments of the *SCL* coding region and the conserved *SCL* poly(A) region all show the expected pattern of sequence conservation (see Table S1 and Fig. S7 at http://hscl.cimr.cam.ac.uk/Gottgens_et_al_MCB_Supplemental_Information.pdf). These data indicate that after

the divergence of birds and mammals, multiple *SCL* *cis*-regulatory elements accrued mutations at an accelerated rate in mammalian ancestors. The same “reversed” pattern of frog/chicken/mammal sequence conservation was identified for a GATA-2 hematopoietic/endothelial enhancer but not for *Runx-1* or *Hex/PRH* (see Table S1 and Fig. S8 at the above URL), suggesting that *cis*-regulatory remodelling coupled with accelerated evolution of regulatory elements may be a more widespread phenomenon. We postulate that this accelerated acquisition of mutations in multiple *cis*-regulatory elements may have been triggered by the creation of a redundant element. The platypus and opossum sequences demonstrate that the switch of the +19 equivalent from a GATA/E-box to an ETS/ETS/GATA motif was an early event. This switch would have resulted in redundancy with the ancestral orthologues of both the promoter and the –4 enhancer, thus freeing other elements to accumulate alterations more rapidly without deleterious consequences. The ancestral *SCL* promoter is active in brain as well as in blood/endothelium, and so there may also be adaptive value in disentangling the individual components of such a hybrid bifunctional element, thus allowing unconstrained evolution of the individual components. We have shown that analogous departures from the expected pattern of regulatory element sequence conservation occur at other loci (e.g., GATA-2) but, depending on the timing of the initial trigger, will not always be revealed by frog/chicken/mammal comparisons.

In summary, our article contains a detailed characterization of the evolution of multiple elements across a key vertebrate regulatory locus, using stringent transgenic functional assays for individual *cis* elements. This analysis revealed remarkable remodeling of *cis*-regulatory elements across the locus despite

FIG. 6. Binding site turnover across widely dispersed regulatory elements (A and B) 6-way sequence alignment of the *SCL* +19 (A) or *SCL* promoter (B) between human, mouse, opossum, platypus, chicken, and frog (shading as for Fig. 2A). Putative transcription factor binding sites are indicated. (C) Luciferase assay of 416B cells stably transfected with the core *SCL* promoter of mouse (200 bp), opossum (200 bp), platypus (250 bp), 5' ETS/ETS/GATA platypus mutant (GATA to GTGC, ETS [TTCC] to TCTC), chicken (187 bp), and chicken 5' deletion (142 bp). A schematic representation of the constructs used is shown on the left. The luciferase assay results are shown on the right as relative luciferase activity increases over results for a promoterless control. Data shown represent the means from at least 4 independent pools of transfected cells. (D) Despite the same conserved expression pattern, the functions of individual *SCL* elements are different in mammals and lower vertebrates. The mammalian promoter has lost ETS sites that have been acquired by the mammalian +19 enhancer. Mammalian +19 has lost the functional E box present in the ancestral locus, but functional GATA/E-box motifs are present in both the mammalian and ancestral +40 enhancers.

constancy of transcriptional inputs and output (i.e., expression pattern). Moreover, an unexpected pattern of sequence conservation in regulatory elements (but not in coding regions) suggests an accelerated accumulation of mutations between 310 million and 200 million years ago, which our results suggest is probably the consequence of the formation of a new but redundant element. Given that redundancy of regulatory elements has been observed for many other vertebrate gene loci, *cis*-regulatory remodeling is likely to be a general phenomenon, suggesting that the comprehensive functional data presented here are of broad relevance for our understanding of vertebrate gene regulation and evolution.

ACKNOWLEDGMENTS

We thank Michelle Hammett, Jacinta Carter, and Sandie Piltz for generating transgenic mice, Amy Chaney and Rebecca Kelley for sectioning embryos, and Pat Simpson and James C. Smith for constructive comments on the manuscript. We are also indebted to Jane Rogers, Darren Grafham, and the entire sequencing teams 40 and 47 at the Wellcome Trust Sanger Institute, Chris Drake for the chicken anti-*SCL* antibody, and Atsushi Miyawaki for the Venus/pSC2 construct.

Work in our laboratories is supported by the Wellcome Trust, the Leukemia Research Fund, the Medical Research Council, and the Leukemia and Lymphoma Society of America.

REFERENCES

- Anand, S., W. C. Wang, D. R. Powell, S. A. Bolanowski, J. Zhang, C. Ledje, A. B. Pawashe, C. T. Amemiya, and C. S. Shashikant. 2003. Divergence of *Hox8* early enhancer parallels diverged axial morphologies between mammals and fishes. *Proc. Natl. Acad. Sci. U. S. A.* **100**:15666–15669.
- Bejerano, G., M. Pheasant, I. Makunin, S. Stephen, W. J. Kent, J. S. Matlack, and D. Haussler. 2004. Ultraconserved elements in the human genome. *Science* **304**:1321–1325.
- Belting, H. G., C. S. Shashikant, and F. H. Ruddle. 1998. Modification of expression and cis-regulation of *Hox8* in the evolution of diverged axial morphology. *Proc. Natl. Acad. Sci. U. S. A.* **95**:2355–2360.
- Bockamp, E. O., F. McLaughlin, A. M. Murrell, B. Gottgens, L. Robb, C. G. Begley, and A. R. Green. 1995. Lineage-restricted regulation of the murine *SCL*/TAL-1 promoter. *Blood* **86**:1502–1514.
- Bradley, C. K., E. A. Takano, M. A. Hall, J. R. Gothert, A. R. Harvey, C. G. Begley, and J. A. van Eekelen. 2006. The essential haematopoietic transcription factor *Scl* is also critical for neuronal development. *Eur. J. Neurosci.* **23**:1677–1689.
- Bronchain, O. J., K. O. Hartley, and E. Amaya. 1999. A gene trap approach in *Xenopus*. *Curr. Biol.* **9**:1195–1198.
- Brown, C. D., D. S. Johnson, and A. Sidow. 2007. Functional architecture and evolution of transcriptional elements that drive gene coexpression. *Science* **317**:1557–1560.
- Caldwell, R. B., P. Fiedler, U. Schoetz, and J. M. Buerstedde. 2007. Gene function analysis using the chicken B-cell line DT40. *Methods Mol. Biol.* **408**:193–210.
- Cameron, R. A., and E. H. Davidson. 2009. Flexibility of transcription factor target site position in conserved cis-regulatory modules. *Dev. Biol.* **336**:122–135.
- Carroll, S. B. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134**:25–36.

11. Chan, W. Y., G. A. Follows, G. Lacaud, J. E. Pimanda, J. R. Landry, S. Kinston, K. Knezevic, S. Piltz, I. J. Donaldson, L. Gambardella, F. Sablitzky, A. R. Green, V. Kouskoff, and B. Gottgens. 2007. The paralogous hematopoietic regulators *Lyl1* and *Scl* are coregulated by Ets and GATA factors, but *Lyl1* cannot rescue the early *Scl*^{-/-} phenotype. *Blood* **109**:1908–1916.
12. Chapman, M. A., F. J. Charchar, S. Kinston, C. P. Bird, D. Grafham, J. Rogers, F. Grutzner, J. A. Graves, A. R. Green, and B. Gottgens. 2003. Comparative and functional analyses of *LYL1* loci establish marsupial sequences as a model for phylogenetic footprinting. *Genomics* **81**:249–259.
13. Chapman, M. A., I. J. Donaldson, J. Gilbert, D. Grafham, J. Rogers, A. R. Green, and B. Gottgens. 2004. Analysis of multiple genomic sequence alignments: a web resource, online tools, and lessons learned from analysis of mammalian *SCL* loci. *Genome Res.* **14**:313–318.
14. Chiu, C. H., C. Amemiya, K. Dewar, C. B. Kim, F. H. Ruddle, and G. P. Wagner. 2002. Molecular evolution of the *HoxA* cluster in the three major gnathostome lineages. *Proc. Natl. Acad. Sci. U. S. A.* **99**:5492–5497.
15. Delabesse, E., S. Ogilvy, M. A. Chapman, S. G. Piltz, B. Gottgens, and A. R. Green. 2005. Transcriptional regulation of the *SCL* locus: identification of an enhancer that targets the primitive erythroid lineage in vivo. *Mol. Cell. Biol.* **25**:5215–5225.
16. Donaldson, I. J., M. Chapman, S. Kinston, J. R. Landry, K. Knezevic, S. Piltz, N. Buckley, A. R. Green, and B. Gottgens. 2005. Genome-wide identification of cis-regulatory sequences controlling blood and endothelial development. *Hum. Mol. Genet.* **14**:595–601.
17. Drake, C. J., S. J. Brandt, T. C. Trusk, and C. D. Little. 1997. *TAL1/SCL* is expressed in endothelial progenitor cells/angioblasts and defines a dorsal-to-ventral gradient of vasculogenesis. *Dev. Biol.* **192**:17–30.
18. Fisher, S., E. A. Grice, R. M. Vinton, S. L. Bessling, and A. S. McCallion. 2006. Conservation of RET regulatory function from human to zebrafish without sequence similarity. *Science* **312**:276–279.
19. Follows, G. A., H. Tagoh, P. Lefevre, D. Hodge, G. J. Morgan, and C. Bonifer. 2003. Epigenetic consequences of AML1-ETO action at the human *c-FMS* locus. *EMBO J.* **22**:2798–2809.
20. Gering, M., A. R. Rodaway, B. Gottgens, R. K. Patient, and A. R. Green. 1998. The *SCL* gene specifies haemangioblast development from early mesoderm. *EMBO J.* **17**:4029–4045.
21. Gompel, N., B. Prud'homme, P. J. Wittkopp, V. A. Kassner, and S. B. Carroll. 2005. Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in *Drosophila*. *Nature* **433**:481–487.
22. Gottgens, B., L. M. Barton, M. A. Chapman, A. M. Sinclair, B. Knudsen, D. Grafham, J. G. Gilbert, J. Rogers, D. R. Bentley, and A. R. Green. 2002. Transcriptional regulation of the stem cell leukemia gene (*SCL*)—comparative analysis of five vertebrate *SCL* loci. *Genome Res.* **12**:749–759.
23. Gottgens, B., L. M. Barton, J. G. Gilbert, A. J. Bench, M. J. Sanchez, S. Bahn, S. Mistry, D. Grafham, A. McMurray, M. Vaudin, E. Amaya, D. R. Bentley, A. R. Green, and A. M. Sinclair. 2000. Analysis of vertebrate *SCL* loci identifies conserved enhancers. *Nat. Biotechnol.* **18**:181–186.
24. Gottgens, B., C. Broccardo, M. J. Sanchez, S. Deveaux, G. Murphy, J. R. Gothert, E. Kotsopoulou, S. Kinston, L. Delaney, S. Piltz, L. M. Barton, K. Knezevic, W. N. Erber, C. G. Begley, J. Frampton, and A. R. Green. 2004. The *scl* +18/19 stem cell enhancer is not required for hematopoiesis: identification of a 5' bifunctional hematopoietic-endothelial enhancer bound by Flt-1 and Elf-1. *Mol. Cell. Biol.* **24**:1870–1883.
25. Gottgens, B., A. Nastos, S. Kinston, S. Piltz, E. C. Delabesse, M. Stanley, M. J. Sanchez, A. Cia-Uitz, R. Patient, and A. R. Green. 2002. Establishing the transcriptional programme for blood: the *SCL* stem cell enhancer is regulated by a multiprotein complex containing Ets and GATA factors. *EMBO J.* **21**:3039–3050.
26. Green, A. R., T. Lints, J. Visvader, R. Harvey, and C. G. Begley. 1992. *SCL* is coexpressed with *GATA-1* in hemopoietic cells but is also expressed in developing brain. *Oncogene* **7**:653–660.
27. Ishibashi, S., K. L. Kroll, and E. Amaya. 2008. A method for generating transgenic frog embryos. *Methods Mol. Biol.* **461**:447–466.
28. Jaffredo, T., K. Bollorot, D. Sugiyama, R. Gautier, and C. Drevon. 2005. Tracing the hemangioblast during embryogenesis: developmental relationships between endothelial and hematopoietic cells. *Int. J. Dev. Biol.* **49**:269–277.
29. Jeong, S., M. Rebeiz, P. Andolfatto, T. Werner, J. True, and S. B. Carroll. 2008. The evolution of gene regulation underlies a morphological difference between two *Drosophila* sister species. *Cell* **132**:783–793.
30. Kontarakis, J., H. H. Chen, A. Riggs, and C. Bonifer. 2000. Chromatin fine structure profiles for a developmentally regulated gene: reorganization of the lysozyme locus before trans-activator binding and gene expression. *Genes Dev.* **14**:2106–2122.
31. Kroll, K. L., and E. Amaya. 1996. Transgenic *Xenopus* embryos from sperm nuclear transplantations reveal FGF signaling requirements during gastrulation. *Development* **122**:3173–3183.
32. Lefevre, P., J. Witham, C. E. Lacroix, P. N. Cockerill, and C. Bonifer. 2008. The LPS-induced transcriptional upregulation of the chicken lysozyme locus involves CTCF eviction and noncoding RNA transcription. *Mol. Cell* **32**:129–139.
33. Ludwig, M. Z., C. Bergman, N. H. Patel, and M. Kreitman. 2000. Evidence for stabilizing selection in a eukaryotic enhancer element. *Nature* **403**:564–567.
34. Ludwig, M. Z., A. Palsson, E. Alekseeva, C. M. Bergman, J. Nathan, and M. Kreitman. 2005. Functional evolution of a cis-regulatory module. *PLoS Biol.* **3**:e93.
35. McGregor, A. P., V. Orgogozo, I. Delon, J. Zanet, D. G. Srinivasan, F. Payre, and D. L. Stern. 2007. Morphological evolution through multiple cis-regulatory mutations at a single gene. *Nature* **448**:587–590.
36. Mead, P. E., C. M. Kelley, P. S. Hahn, O. Piedad, and L. I. Zon. 1998. *SCL* specifies hematopoietic mesoderm in *Xenopus* embryos. *Development* **125**:2611–2620.
37. Mikkelsen, T. S., M. J. Wakefield, B. Aken, C. T. Amemiya, J. L. Chang, S. Duke, M. Garber, A. J. Gentles, L. Goodstadt, A. Heger, J. Jurka, M. Kamal, E. Mauceli, S. M. Searle, T. Sharpe, M. L. Baker, M. A. Batzer, P. V. Benos, K. Belov, M. Clamp, A. Cook, J. Cuff, R. Das, L. Davidow, J. E. Deakin, M. J. Fazzari, J. L. Glass, M. Grabherr, J. M. Greally, W. Gu, T. A. Hore, G. A. Huttley, M. Kleber, R. L. Jirtle, E. Koira, J. T. Lee, S. Mahony, M. A. Marra, R. D. Miller, R. D. Nicholls, M. Oda, A. T. Papenfuss, Z. E. Parra, D. D. Pollock, D. A. Ray, J. E. Schein, T. P. Speed, K. Thompson, J. L. VandeBerg, C. M. Wade, J. A. Walker, P. D. Waters, C. Webber, J. R. Weidman, X. Xie, M. C. Zody, J. A. Graves, C. P. Ponting, M. Breen, P. B. Samollow, E. S. Lander, and K. Lindblad-Toh. 2007. Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. *Nature* **447**:167–177.
38. Nagai, T., K. Ibata, E. S. Park, M. Kubota, K. Mikoshiba, and A. Miyawaki. 2002. A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat. Biotechnol.* **20**:87–90.
39. Ogilvy, S., R. Ferreira, S. G. Piltz, J. M. Bowen, B. Gottgens, and A. R. Green. 2007. The *SCL* +40 enhancer targets the midbrain together with primitive and definitive hematopoiesis and is regulated by *SCL* and *GATA* proteins. *Mol. Cell. Biol.* **27**:7206–7219.
40. Porcher, C., W. Swat, K. Rockwell, Y. Fujiwara, F. W. Alt, and S. H. Orkin. 1996. The T cell leukemia oncoprotein *SCL/tal-1* is essential for development of all hematopoietic lineages. *Cell* **86**:47–57.
41. Prud'homme, B., N. Gompel, and S. B. Carroll. 2007. Emerging principles of regulatory evolution. *Proc. Natl. Acad. Sci. U. S. A.* **104**(Suppl.):8605–8612.
42. Prud'homme, B., N. Gompel, A. Rokas, V. A. Kassner, T. M. Williams, S. D. Yeh, J. R. True, and S. B. Carroll. 2006. Repeated morphological evolution through cis-regulatory changes in a pleiotropic gene. *Nature* **440**:1050–1053.
43. Robb, L., N. J. Elwood, A. G. Elefanti, F. Kontgen, R. Li, L. D. Barnett, and C. G. Begley. 1996. The *scl* gene product is required for the generation of all hematopoietic lineages in the adult mouse. *EMBO J.* **15**:4123–4129.
44. Sanchez, M., B. Gottgens, A. M. Sinclair, M. Stanley, C. G. Begley, S. Hunter, and A. R. Green. 1999. An *SCL* 3' enhancer targets developing endothelium together with embryonic and adult haematopoietic progenitors. *Development* **126**:3891–3904.
45. Silberstein, L., M. J. Sanchez, M. Socolovsky, Y. Liu, G. Hoffman, S. Kinston, S. Piltz, M. Bowen, L. Gambardella, A. R. Green, and B. Gottgens. 2005. Transgenic analysis of the stem cell leukemia +19 stem cell enhancer in adult and embryonic hematopoietic and endothelial cells. *Stem Cells* **23**:1378–1388.
46. Sinclair, A. M., B. Gottgens, L. M. Barton, M. L. Stanley, L. Pardanaud, M. Klaine, M. Gering, S. Bahn, M. Sanchez, A. J. Bench, J. L. Fordham, E. Bockamp, and A. R. Green. 1999. Distinct 5' *SCL* enhancers direct transcription to developing brain, spinal cord, and endothelium: neural expression is mediated by *GATA* factor binding sites. *Dev. Biol.* **209**:128–142.
47. Sive, H. L., R. M. Grainger, and R. M. Harland. 2000. Early development of *Xenopus laevis*: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
48. Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**:4673–4680.
49. Tumpel, S., M. Maconochie, L. M. Wiedemann, and R. Krumlauf. 2002. Conservation and diversity in the cis-regulatory networks that integrate information controlling expression of *Hoxa2* in hindbrain and cranial neural crest cells in vertebrates. *Dev. Biol.* **246**:45–56.
50. Wadman, I. A., H. Osada, G. G. Grutz, A. D. Agulnick, H. Westphal, A. Forster, and T. H. Rabbitts. 1997. The LIM-only protein *Lmo2* is a bridging molecule assembling an erythroid, DNA-binding complex which includes the *TAL1*, *E47*, *GATA-1* and *Ldb1/NLI* proteins. *EMBO J.* **16**:3145–3157.
51. Zhang, X. Y., and A. R. Rodaway. 2007. *SCL*-GFP transgenic zebrafish: in vivo imaging of blood and endothelial development and identification of the initial site of definitive hematopoiesis. *Dev. Biol.* **307**:179–194.