Haematopoiesis represents one of the best studied models of adult stem cell development and differentiation [1,2]. Powerful techniques allow purification and in vitro as well as in vivo functional assays of small subsets of cells, from haematopoietic stem cells (HSCs) via a plethora of intermediate progenitors to fully mature cell types. Transcription factors (TFs) directly regulate gene expression and thus control cellular phenotypes. It is no surprise, therefore, that TFs have emerged as some of the most powerful regulators of both normal development and disease.

TFs play important roles during haematopoiesis, from stem cell maintenance to lineage commitment and differentiation. However, relatively little is known about the way in which regulatory information is encoded in the genome, and how individual TFs are integrated into wider regulatory networks. Based on the recent analysis of large-scale efforts to reconstruct tissue-specific regulatory networks, it has been suggested that transcriptional regulatory networks are characterised by a high degree of connectivity between TFs and transcriptional cofactors. Extensive cross- and autoregulatory links therefore create densely connected regulatory circuits that control the large numbers of tissue-specific effector proteins (enzymes, structural proteins) [3,4] (Figure 1). To understand the functionality of large mammalian regulatory networks, it will therefore be important to identify downstream target genes of specific TFs as well as gain insight into combinatorial TF interactions. This in turn will not only provide fundamental insights into normal development, but also advance our understanding of how deregulation of networks contributes to pathology.

The cis-regulatory regions of a gene locus can be thought of as different modules, each partaking in an important role, such as driving expression of the gene to a specific subset of cells or a specific tissue type. The activity of each regulatory region is controlled by a distinct set of upstream regulators. The individual regulatory regions within a given gene locus may have overlapping or very distinct upstream regulators, and it is the combined activity of all these regions that ultimately controls gene expression. Comprehensive identification and characterisation of true functional cis-regulatory regions therefore represent an essential prerequisite to integrate important regulatory genes into wider transcriptional networks. Traditionally, DNaseI mapping was performed to identify regions of open/accessible chromatin. More recently, comparative genomic sequence analysis has been used to identify highly conserved sequences, which were taken to represent candidate regulatory elements based on the premise that sequence conservation indicated an important function [5-7]. The most recent development has been that of whole genome re-sequencing, which when coupled with chromatin...
immunoprecipitation assays allows genome-wide mapping of the chromatin status for a given histone modification [8]. Though more predictive than previous approaches, these techniques still require functional validation of candidate elements, which involves in vivo and in vitro experiments to assess the true function of a given candidate regulatory region.

Several gene loci coding for TFs essential for haematopoiesis have been characterised using a combination of the above techniques. Collectively, these studies provided important insights into TF hierarchies and regulatory network core circuits [9-11]. This review will specifically focus on three haematopoietic loci, encoding the key haematopoietic regulators Scl/Tal1, Lmo2 and Gfi1.

Transcriptional regulation of Scl
The basic helix-loop-helix TF Scl/Tal1 is a key regulator of haematopoiesis with additional important roles in the development of the vascular and central nervous systems [12-16]. Within the haematopoietic system, Scl is essential for the development of HSCs as well as further differentiation into the erythroid and megakaryocytic lineages [17].

Since correct spatio-temporal expression of Scl is crucial for the appropriate execution of its biological functions, much effort has been invested into understanding how Scl is regulated. Using a combination of long-range comparative sequence analysis and both in vitro and in vivo functional analysis, multiple cis-regulatory elements have been identified in the murine Scl locus, each of which directs expression to a subdomain of endogenous Scl expression when tested in transgenic mice (Figure 2). Scl has three promoters located in different exons (exons 1a, 1b and exon 4), none of which displayed haematopoietic activity when tested in transgenic mice. A search for additional cis-regulatory elements led to the identification of three haematopoietic enhancers (-4, +19 and +40 kb). The -4 Scl enhancer, characterized by the presence of five Ets sites, drives expression to endothelium and fetal blood progenitors [18]. The +19 Scl enhancer was shown to drive expression of Scl in HSCs, haematopoietic progenitors and endothelial cells [19-21] and critically depended on an Ets/Ets/GATA composite motif shown to be bound in vivo by Elf-1, Fli-1 and Gata2 [22]. Of note, the +19 enhancer was flanked by a nearby hypersensitive site (+18 Scl element), which did not function as an enhancer but contains a mammalian interspersed repeat that is essential for its ability to ‘boost’ activity of the +19 element [23]. The +40 Scl enhancer drives expression to erythroid cells [24,25] as well as midbrain and is characterized by the presence of two Gata/E-box motifs. Mutation or deletion of a single one of these motifs leads to a loss of function of the enhancer [24,25].

Taken together, these studies have highlighted the presence of three haematopoietic enhancers within the murine Scl locus, with distinct yet overlapping regulatory codes that contribute to the overall correct spatio-temporal expression of Scl. Interestingly, a recent study comparing the functionality of the mouse Scl enhancers with their corresponding chicken counterparts suggested that elements shared by mammals and lower vertebrates exhibit functional differences and binding site turnover between widely separated cis-regulatory modules [26]. Remarkably, however, the regulatory inputs and overall expression patterns remain the same across different species. This in turn suggested that significant regulatory changes may be widespread, and not only apply to genes with altered expression patterns, but also to those where expression is highly conserved.

Transcriptional regulation of Lmo2
The Lim domain only 2 gene (Lmo2) encodes a transcriptional cofactor that is essential for haematopoiesis [27,28]. The Lmo2 protein does not bind to DNA directly but rather participates in the formation of multipartite DNA-binding complexes with other TFs, such as Ldb1, Scl/Tal1, E2A and Gata1 or Gata2 [29-31]. Lmo2 is widely expressed across haematopoiesis with the exception of mature T-lymphoid cells where aberrant expression of Lmo2 results in T-cell leukaemias [32].

Lmo2 contains three promoters: the proximal promoter, which drives the majority of expression in endothelial cells [33]; the distal promoter, which is active in the fetal liver and specific T-cell acute lymphoblastic leukemia (T-ALL) cell lines [34]; and the intermediate promoter, which was detected in CD34+ cells and was implicated in mediating LMO2 expression in T-ALL patients where high levels of LMO2 were present in the absence of any...
translocation involving the LMO2 locus [35]. However, none of the three promoters on their own displayed robust expression when tested in transgenic mice [33,36], which led to the identification of eight enhancer elements dispersed over 100 kb that could recapitulate the expression of Lmo2 in normal haematopoiesis [36]. Of note, while individual elements augmented endothelial expression of the proximal promoter, robust haematopoietic expression was only observed when they were combined together (Figure 3). This type of combinatorial collaboration between regulatory elements to obtain haematopoietic activity has been seen for other gene loci, such as Endoglin [37], suggesting a process of step-wise and modular activation of the locus during the development of blood and endothelial cells from their common precursor.

Transcriptional regulation of Gfi1
The Growth factor independence 1 gene (Gfi1) was originally identified in a retroviral screen designed to identify regulatory pathways that could initiate interleukin-2 independence in T cells [38]. Within the haematopoietic system Gfi1 is expressed in HSCs [39], specific subsets of T cells [40], granulocytes, monocytes, and activated macrophages [41]. Gfi1+/− mice lack neutrophils [41,42] and Gfi1−/− HSCs are unable to maintain long-term haematopoiesis because elevated levels of proliferation lead to eventual exhaustion of the stem cell pool [39,43]. Outside the haematopoietic system, Gfi1 is also specifically expressed in sensory epithelia, the lungs, neuronal precursors, the inner ear, intestinal epithelia and during mammary gland development [44-47].

A recent study used a combination of comparative genomics, locus-wide chromatin immunoprecipitation assays and functional validation within cell lines and transgenic animals to identify cis-regulatory regions within the Gfi1 locus [48]. Four regulatory regions (-3.4 kb min pro, -1.2 kb min pro, +5.8 kb enhancer and +35 kb enhancer) were shown to recapitulate endogenous expression patterns of Gfi1 in the central nervous system,
gut, limbs and developing mammary glands but no haematopoietic staining was observed. However, a recent genome-wide ChIP-Seq experiment [49] revealed binding of Scl/Tal1 to a region situated 35 kb upstream of the Gfi1 promoter within the last intron of its 5’ flanking gene, Evi5. This element was subsequently validated in transgenic assays, which demonstrated lacZ staining at multiple sites of haematopoietic stem/progenitor cell emergence (vitelline vessels, fetal liver, and dorsal aorta). Moreover, the element was also shown to be bound by TFs known to be critical for haematopoiesis, including Scl/Tal1, Pu.1/Sfpi1, Runx1, Erg, Meis1, and Gata2, thus integrating Gfi1 into the wider HSC regulatory network. This study therefore supports the notion that important regulatory elements can be located at a significant distance from the gene they control (Figure 4), and thus emphasize the need for careful interpretation of genome-wide TF binding datasets [49,50].

Figure 3. Combinatorial interactions of distinct enhancers are critical to recapitulate the endogenous expression of Lmo2. (a) The Lmo2 gene locus is drawn to scale. Exons are shown as black rectangles. Regulatory elements (-75/-70/-25/-12/pP/+1) are highlighted using shapes and distinct colours (-75 = orange diamond; -70 = green octagon; -25 = blue oval; pP = red rectangle; -12 = red triangle; +1 = purple triangle). (b) Transgenic animals were generated with many different combinations of the identified regulatory elements. The -75 enhancer and pP showed strong expression in endothelium, circulating erythrocytes and fetal liver. The -70 enhancer together with pP showed weak staining in endothelium and haematopoietic progenitor cells. The -25 or the -12 enhancer together with pP showed weak staining in endothelium and haematopoietic progenitor cells. The +1 enhancer with pP gave rise to lacZ staining in the tail, apical ridge of the limbs, fetal liver and strong endothelium. Only when these elements were coupled together was a staining pattern corresponding to endogenous expression of Lmo2 seen [36]. Strength of staining is indicated: +++, very strong; ++, intermediate; +, weak; -, not present.
Figure 4. Combinatorial transcription factor binding identified the Gfi1 –35 kb regulatory region. Raw ChIP-Seq read data from [50] were transformed into a density plot for each transcription factor and loaded into the UCSC genome browser as custom tracks above the UCSC tracks for gene structure and mammalian homology. A discreet binding event for all ten TFs (Scl/Tal1, Lyl1, Lmo2, Gata2, Runx1, Meis1, Pu.1, Fli1, Erg and Gfi1b) can be seen in the last intron of the 5' flanking gene, Evi5 (indicated by an asterisk). This region was subsequently shown to drive expression in early haematopoietic cells in transgenic mouse embryos [48].
Transcriptional regulation of other key haematopoietic transcription factors

The transcriptional control of several other TFs known to play important roles within haematopoiesis have also been investigated. Runx1 has been shown to be transcribed from two promoter elements, both of which collaborate with the Runx1 +23 kb enhancer to drive expression of Runx1 to sites of HSC emergence [51-53]. Moreover, the Runx1 +23 kb region was shown to be regulated by important haematopoietic TFs (Gata2, Fli1, Elf1, Pu.1, Scl, Lmo2, Ldb1 and Runx1 itself) [53,54]. Lyl1 is known to contain a promoter region that can be divided into two separate promoter elements that are responsible for driving the expression of Lyl1 within endothelial, haematopoietic progenitor, and megakaryocytic cells [55]. These promoter elements were shown to contain conserved Ets and Gata motifs that were bound in vivo by Fli1, Elf1, Erg, Pu.1, and Gata2. Multiple elements within the Gata2 locus have been identified (-77 kb, -3.9 kb, -2.8 kb, -1.8 kb, +9.5 kb and 1s promoter) [56-58] with the -1.8 kb region being essential for maintaining Gata2 repression in terminally differentiating cells [58]. Elf1 contains four promoter elements (-55 kb, -49 kb, -21 kb and proximal), which are used in a cell-type-specific manner in combination with a lineage-specific -14 kb enhancer element [59]. Enhancer elements utilising the Ets/Ets/Gata regulatory code, originally defined in the Scl +19 enhancer, were also identified in the Fli1, Gata2, Hhex/Prh and Smad6 gene loci [5,57]. The picture emerging, therefore, is that transcriptional control of important haematopoietic TF loic is achieved through multiple regulatory elements but the number of upstream regulators may be relatively small. The same binding motifs are repeatedly found, but it is the precise arrangement within a single element as well as the interactions between elements that ultimately control expression.

Conclusion

Recent analysis of gene regulatory networks controlling pluripotency in embryonic stem cells suggests that a finite number of major combinatorial interactions are critical in controlling cellular phenotypes [60,61]. Identification and subsequent functional characterisation of specific regulatory elements provides a powerful route into deciphering these combinatorial regulatory interactions. Whilst traditional methods of identifying regulatory elements should not be overlooked, it is essential to integrate new genome-wide methods to ensure that regulatory elements outside traditional gene loci boundaries are not overlooked. With the genome-wide mapping of TF binding events now eminently feasible, the importance of sequence conservation as a primary technique for identification of regulatory elements will diminish. Nevertheless, genome-wide mapping of binding events is descriptive and therefore no substitute for conventional functional assays, which are therefore likely to remain an important component of any research programme aimed at elucidating transcriptional control mechanisms.

References

1. Orkin SH, Zon LI: SnapShot: hematopoiesis. Cell 2008, 132:712.
2. Ottersbach K, Smith A, Wood A, Gottgens B: Ontogeny of haematopoiesis: recent advances and open questions. Br J Haematol 2010, 148:343-355.
3. Davidson EH: Emerging properties of animal gene regulatory networks. Nature 2010, 468:911-920.
4. modENCODE Consortium, Roy S, Ernst J, Kharchenko PV, Kherdipour P, Negre N, Eaton ML, Landolin JM, Bristow CA, Ma L, Lin MF, Washierl S, Arshinoff BI, Ay F, Meyer PE, Robine N, Washington NL, Di Stefano L, Berezikov E, Brown CD, Candeaia R, Carlson JW, Carr A, Jungreis I, Marbach D, Sealfon R, Tolstorukov MY, Will S, Alekseyenko AA, Arteri C, et al: Identification of functional elements and regulatory circuits by Drosophila modENCODE. Science 2010, 330:1787-1797.
5. Donaldson JJ, Chapman M, Kinston S, Landry JR, Knezevic K, Pitts S, Buckley N, Green AR, Gottgens B: Genome-wide identification of cis-regulatory sequences controlling blood and endothelial development. Hum Mol Genet 2005, 14:595-611.
6. Gottgens B, Barton LM, Gilbert JG, Bench AJ, Sanchez MJ, Bahn S, Misty S, Graffham D, McMurray A, Vaudin M, Arshinoff BI, Bentley DR, Green AR, Sinclair AM: Analysis of vertebrate SCL loci identifies conserved enhancers. Nat Biotechnol 2000, 18:181-186.
7. Loots GG, Locksley RM, Blakespoor CM, Wang ZE, Miller W, Rubin EM, Frazer KA: Identification of a coordinate regulator of interleukins 4, 13, and 5 by cross-species sequence comparisons. Science 2000, 288:136-140.
8. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K: High-resolution profiling of histone methylations in the human genome. Cell 2007, 129:823-837.
9. Foster SD, Osman SH, Wilson NK, Gottgens B: From genes to cells to tissues - modelling the haematopoietic system. Mol Biosyst 2009, 5:1413-1420.
10. Miranda-Saavedra D, Gottgens B: Transcriptional regulatory networks in haematopoiesis. Curr Opin Genet Dev 2008, 18:530-535.
11. Pimanda JE, Gottgens B: Gene regulatory networks governing haematopoietic stem cell development and identity. Int J Dev Biol 2010, 54:1201-1211.
12. Kallianpur AR, Jordan JE, Brandt SJ: The SCL/TAL-1 gene is expressed in progenitors of both the hematopoietic and vascular systems during embryogenesis. Blood 1994, 83:1200-1208.
13. Muroyama Y, Fujwara Y, Orkin SH, Rowitch DH: Specification of astrocytes by bHLH protein SCL in a restricted region of the neural tube. Nature 2005, 438:360-363.
14. Visvader JE, Fujwara Y, Orkin SH: Unsuspected role for the T-cell leukemia protein SCL/tal-1 in vascular development. Genes Dev 1998, 12:473-479.
52. Bee T, Liddiard K, Swiers G, Bickley SR, Vink CS, Jarratt A, Hughes JR, Medvinsky A, de Bruijn MF: Alternative Runx1 promoter usage in mouse developmental hematopoiesis. Blood Cells Mol Dis 2009, 43:35-42.

53. Nottingham WT, Jarratt A, Burgess M, Speck CL, Cheng JF, Prabhakar S, Rubin EM, Li PS, Sloane-Stanley J, Kong ASJ, de Bruijn MF: Runx1-mediated hematopoietic stem-cell emergence is controlled by a Gata/Ets/SCL-regulated enhancer. Blood 2007, 110:4188-4197.

54. Landry JR, Kinston S, Knezevic K, de Bruijn MF, Wilson N, Nottingham WT, Petitz M, Edenhofer F, Pimanda JE, Ottersbach K, Gottgens B: Runx genes are direct targets of Scl/Tal1 in the yolk sac and fetal liver. Blood 2008, 111:3005-3014.

55. Chan WY, Follows GA, Lacaud G, Pimanda JE, Landry JR, Kinston S, Knezevic K, Piltz S, Donaldson U, Gambardella L, Sablitzky F, Green AR, Kouskoff V, Gottgens B: The paralogous hematopoietic regulators Lyf1 and Scl are coregulated by Ets and GATA factors, but Lyf1 cannot rescue the early Scl-/- phenotype. Blood 2007, 109:1908-1916.

56. Grass JA, Jing H, Kim SI, Martowicz ML, Pal S, Blobel GA, Bresnick EH: Distinct functions of dispersed GATA factor complexes at an endogenous gene locus. Mol Cell Biol 2006, 26:7056-7067.

57. Pimanda JE, Ottersbach K, Knezevic K, Kinston S, Chan WY, Wilson NK, Landry JR, Wood AD, Kolb-Kokocinski A, Green AR, Tannahill D, Lacaud G, Kouskoff V, Gottgens B: Gata2, Flt1, and Scl form a recursively wired gene-regulatory circuit during early hematopoietic development. Proc Natl Acad Sci U S A 2007, 104:17692-17697.

58. Snow JW, Trowbridge JJ, Fujiwara T, Emambokus NE, Grass JA, Orkin SH, Bresnick EH: A single cis element maintains repression of the key developmental regulator Gata2. PLoS Genet 2010, 6:e1001103.

59. Calero-Nieto FJ, Wood AD, Wilson NK, Kinston S, Landry JR, Gottgens B: Transcriptional regulation of Elf-1: locus-wide analysis reveals four distinct promoters, a tissue-specific enhancer, control by PU.1 and the importance of Elf-1 downregulation for erythroid maturation. Nucleic Acids Res 2010, 38:6363-6374.

60. Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, Wong E, Orlov YL, Zhang W, Jiang J, Loh YH, Yeo HC, Yeo ZX, Narang V, Govindarajan KR, Leong B, Shahab A, Ruan Y, Bourque G, Sung WK, Clarke ND, Wei CL, Ng HH: Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell 2008, 133:1106-1117.

61. Kim J, Chu J, Shen X, Wang J, Orkin SH: An extended transcriptional network for pluripotency of embryonic stem cells. Cell 2008, 132:1049-1061.

doi:10.1186/scrt47

Cite this article as: Wilson NK, et al: Transcriptional regulation of hematopoietic transcription factors. Stem Cell Research & Therapy 2011, 2:6.