The SCL Gene
and
Transcriptional Control of Haematopoiesis

Pawandeep Dhami
Darwin College

Thesis submitted for the degree of
Doctor of Philosophy

University of Cambridge
2005

Disclaimer

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specifically indicated in the text.

This Dissertation does not exceed the word limit prescribed by the Biology Degree Committee

Pawandeep Dhami

Wednesday, 30 November 2005

This thesis is dedicated to my mum and dad

Abstract

The SCL Gene and Transcriptional Control of Haematopoiesis

Understanding the events which occur as stem cells differentiate into committed cell lineages is a fundamental issue in cell biology. It has been shown that the SCL transcription factor, also known as TAL1, is central to the mechanisms whereby pluripotent stem cells differentiate into haematopoietic stem cells (HSCs) that ultimately give rise to the various blood lineages. While this process is thought to be tightly regulated at the level of gene expression, the exact ways in which SCL helps direct this process is not well understood. To further understand the biology of SCL and the key regulatory interactions it is involved in during blood development, the powerful techniques of genomic microarray resources in combination with chromatin immunoprecipitation (ChIP-chip) were used. High resolution (400-500 bp) genomic tiling path microarrays spanning the human and mouse SCL loci were constructed. ChIP-chip experiments using a large battery of antibodies raised against various histone modifications, transcription factors, and other regulatory proteins were performed in a number of SCL expressing and non-expressing cell lines. The ChIP material used in the array experiments was not amplified prior to hybridization and resulted in ChIP-chip assays which were reproducible, robust and as sensitive as real-time PCR. Based on the ChIP-chip data that was generated, relationships between transcriptional regulatory events and the underlying DNA sequence were studied across the SCL locus. The results described in this thesis will greatly accelerate our understanding of important biological events which are essential for the expression of SCL, as well as provide insights into mechanisms of mammalian gene regulation likely to be widely applicable.

Acknowledgements

I would like to thank my supervisor Dave Vetrie for all the help, guidance, advice and support provided throughout the course and most importantly being my biggest source of inspiration, especially during difficult times.

I would also like to thank our team of collaborators at the Department of Haematology, University of Cambridge – Tony Green and Bertie Gottgens for their help and advice during the course of this study. Thanks to Eric Delabesse who taught me how to do ChIP, and Ian Donaldson's help in performing the sequence analysis is greatly appreciated.

A number of people at The Sanger Institute have made invaluable contributions to the work presented in this thesis. Special thanks to Philippe Couttet for help with the real-time PCR and expression analysis, Jonathan Cooper for doing all the enhancer trap experiments, and Shane Dillon for performing the CTCF ChIP experiment. Rob Andrews provided great help with designing the primers for the SCL tiling array and Oliver Dovey for printing such beautiful arrays. Heike Fiegler and Philippa Carr helped with learning how to perform array hybridisations. Thanks to Rebecca Curley and Susan Gribble for performing the array-CGH experiments, Hazel Arbery and Ruth Bennett for the DNA extractions and Bee Ling Ng for flow-sorting.

Special thanks to all the support teams at The Sanger Institute particularly, the Media preparation team, the glassware team, the Library team, the Sequencing team, and the Systems support team.

Thanks to the three voices of sanity - Sheila, Cords and Susan, for keeping my spirits up over endless cups of tea and gossip.

My friends Anu and Bharti deserve special thanks. Anu for keeping the telephone lines across the Atlantic very busy and listening to my woes at all times. Bharti has been a constant support for the past one year and shared my highs and lows.

A big thanks to my family in India, who despite being so far away, showered me with lots of love, encouragement and prayers for me to finish this successfully. They were keen for me to finish this just so that I can spend more time with them!

I do not have words to thank you, Arvind. I could not have done this without you.

Table of Contents

Chapter 1 Introduction

1.1. Coding regions of the human genome 1
1.2. Non-coding regions of the human genome 2
1.3. Understanding the regulation of gene expression 3
1.4. Types of non-coding regulatory elements 3
    1.4.1. Promoters 4
    1.4.2. Enhancers 5
    1.4.3. Silencers or repressors 6
    1.4.4. Insulators 6
    1.4.5. Locus control regions (LCRs) 6
    1.4.6. Scaffold/Matrix attachment regions (S/MARs) 6
1.5. Proteins involved in transcriptional regulation 6
    1.5.1. RNA polymerase and basal transcription factors 7
    1.5.2. Sequence-specific transcription factors 9
    1.5.3. Proteins involved in chromatin remodelling 11
1.6. Identification of genomic non-coding regulatory elements 14
    1.6.1. Classical methods 15
    1.6.2. Computational methods 18
    1.6.3. Microarray-based methods 20
    1.6.4. Applications of DNA microarrays 21
    1.6.5. Other ChIP-related methods 26
1.7. Chromatin, epigenetics and transcriptional regulation 26
    1.7.1. Nucleosome: the structural and functional subunit of chromatin 26
    1.7.2. Histone modifications 28
    1.7.3. The histone code 33
1.8. Haematopoiesis 34
1.9. The SCL Gene 37
    1.9.1. SCL gene structure 37
    1.9.2. SCL gene expression 38
    1.9.3. The SCL protein product 41
    1.9.4. SCL function 42
    1.9.5. Regulation of SCL 49
    1.9.6. Physiological downstream targets of SCL 52
1.10. Aims of this thesis 53

Chapter 2 Materials and Methods

2.1. Composition of Solutions 55
2.2. Reagents 58
2.3. Cells and Cell lines 59

2.4. Bacterial clones 59
2.5. Tissue Culture 60
  2.5.1. Culturing of all Cell lines 60
  2.5.2. Cell cryopreservation 61
2.6. DNA Preparation 61
  2.6.1. Preparation of Total Genomic DNA........ 61
  2.6.2. Preparation of DNA from BAC and PAC clones 62
  2.6.3. Chromatin Immunoprecipitation (ChIP) 63
2.7. Construction of the SCL Genomic Tiling Path Microarray 67
  2.7.1. Generation of Human and Mouse SCL Genomic Tiling Path Amplicons 67
  2.7.2. Microarray Slide Printing and Processing 69
2.8. Hybridisation of the array 70
  2.8.1. Random Labeling of DNA samples 70
  2.8.2. Purification of labeled DNA Samples 71
  2.8.3. Manual set-up of array hybridisations 72
  2.8.4. Array hybridisation set-up using the Tecan hybridisation station 74
  2.8.5. Scanning and Data Analysis for ChIP-chip experiments 76
2.9. Sequence Analysis 77
2.10. Statistical analysis 78
2.11. Cluster analysis of histone modification data 78
2.12. Real-time PCR 79
2.13. Procedures for fluorescence in-situ hybridisation (FISH) 80
  2.13.1. Preparation of metaphase chromosomes 80
  2.13.2. Probe labelling by nick translation 81
  2.13.3. Metaphase spread slide preparation 81
  2.13.4. Hybridisation of metaphase spreads 82
  2.13.5. Detection of labeled probes 83
  2.13.6. Acquisition of FISH images 84
2.14. RNA Extraction 84

Chapter 3 Construction and Validation of the SCL Genomic Tiling Path Microarrays and Characterization of SCL Expressing and Non-expressing Cell Lines

3.1. Introduction........ 86
3.2. Aims of this chapter 88
3.3. Construction of the SCL Genomic Tiling Path Microarray 89
  3.3.1. The array chemistry 89
  3.3.2. Defining the genomic regions to be represented on the SCL arrays 90
  3.3.3. Primer design and PCR amplification 90
  3.3.4. Controls included on the arrays 92
3.4. Validation of the SCL genomic tiling path microarrays 93
  3.4.1. Assessment of human and mouse SCL array amplicons by electrophoresis 93

  - 3.4.2. Sequence analysis of the SCL amplicons 94
  - 3.4.3. Validation of the human array elements in genomic array assays 94
- 3.5. Characterization of human SCL expressing and non-expressing cell lines 97
  - 3.5.1. Real-time PCR analysis of the SCL expression in the four cell lines 98
  - 3.5.2. 1 Mb genomic microarray-CGH analysis of K562, Jurkat, HL60 and HPB-ALL 99
  - 3.5.3. Array-CGH analysis of K562, Jurkat, HL60 and HPB-ALL using the SCL genomic tiling path array 103
  - 3.5.4. FISH analysis of K562, Jurkat, HL60 and HPB-ALL with two DNA clones spanning the human SCL locus 105
- 3.6. Discussion 108
  - 3.6.1. SCL genomic tiling path array: sensitivity and resolution issues 108
  - 3.6.2. Characterizing human haematopoietic cell lines 110
  - 3.6.3. Conclusions 112

Chapter 4 Assessment of the SCL Tiling Arrays for use in ChIP-chip Analysis

- 4.1 Introduction 113
- 4.2 Aims of the chapter 115
- 4.3 Overall Strategy 116
- 4.4 Establishing criteria for performing chromatin immunoprecipitation and microarray hybridisations 116
- 4.5 Performance of the SCL genomic tiling path arrays in detecting regions of histone H3 K9/14 diacetylation in human and mouse 119
  - 4.5.1 Assessing the performance of SCL array in ChIP-chip assays for H3 K9/14 diacetylation in K562 cell line 119
  - 4.5.2 Histone H3 acetylation coincides with regulatory function at the SCL locus in human 121
  - 4.5.3 Histone H3 acetylation array data correlates with real-time PCR of H3 acetylation ChIP material 122
  - 4.5.4 ChIP-chip profiles for Histone H3 K9/14 diacetylation in 416B cell line 123
  - 4.5.5 Sequence conservation at sites of histone H3 acetylation in human and mouse 125
- 4.6 Performance of the human SCL genomic tiling path array in detecting binding sites for the GATA-1 transcription factor in K562 cells 127
  - 4.6.1 Assessing reproducibility of the array elements in ChIP-chip assay for GATA-1 transcription factor 127
  - 4.6.2 Identification of regions enriched for GATA-1 across the SCL locus in K562 cell line 129
  - 4.6.3 GATA-1 array data correlates with real-time PCR of GATA-1 ChIP material 130
  - 4.6.4 Conserved sequence analysis of genomic regions showing enrichments with GATA-1 ChIP in K562 131
- 4.7 Performance of the human SCL genomic tiling array in detecting variations in nucleosome density in K562 cells 133
  - 4.7.1 Assessing the reproducibility of the array elements in ChIP-chip assay for histone H3 133

4.7.2 Analysis of nucleosome density at the human SCL locus 135
4.7.3 Normalization using nucleosome density 136
4.8 Discussion 137
4.8.1 Histone H3 Acetylation at the SCL Locus 138
4.8.2 GATA-1 and SCL Regulation 139
4.8.3 Nucleosome Density at the SCL Locus 140
4.8.4 Conclusions 141

Chapter 5 Application of ChIP-chip to elucidate a Comprehensive Set of Regulatory Interactions at the SCL locus

5.1 Introduction 142
5.2 Aims of this chapter 144
5.3 Overall strategy 145
5.4 Identifying and testing antibodies to use in ChIP-chip assays 146
5.5 Constructing maps of *in vivo* DNA-protein interactions at the SCL locus 150
5.6 Histone H3 acetylation patterns at the SCL locus in SCL expressing and non-expressing cell lines 150
5.6.1 Histone H3 K9/14 diacetylation correlates with active genes in human cell lines 150
5.6.2 Histone H3 K9/14 diacetylation is present at SCL regulatory regions in mouse ES cells which do not express SCL 152
5.6.3 Histone H3 modifications at lysines 9, 14, 18 and 27 correlate with transcriptional activity of genes 154
5.7 Histones H2A and H2B acetylation patterns at the SCL locus in SCL expressing and non-expressing cell lines 157
5.7.1 Histone H2A acetylation mimics that of H3 K9/14 diacetylation 157
5.7.2 Acetylation of histone H2B marks tissue-specific genes 157
5.8 Histone H4 acetylation patterns at the SCL locus in SCL expressing and non-expressing cell lines 158
5.8.1 H4 acetylation coincides with H3 acetylation across the SCL locus 158
5.8.2 Acetylation of histone H4 at specific lysine residues correlates variably with transcriptional activity at the SCL locus 161
5.9 Histone H3 lysine methylation patterns at the SCL locus in SCL expressing and non-expressing human and mouse cell-lines 162
5.9.1 The 5’ ends of active genes are hypomethylated for histone H3 lysine 4 monomethylation (H3 K4Me1) 163
5.9.2 Dimethylation and trimethylation at lysine 4 of histone H3 (H3 K4Me2, H3 K4Me3) occur at transcriptionally active genes across the SCL locus in human and mouse 168
5.9.3 Trimethylation of histone H3 lysine 79 (H3 K79Me3) marks the immediate 3’ downstream regions of promoters of active genes 172
5.10 Histone H3 phosphorylation at the SCL locus in SCL expressing and non-expressing human cell lines 173

5.11 Analysis of nucleosome density at the human SCL locus in Jurkat, HL60 and HPB-ALL cell lines 174
5.12 SCL itself binds to the erythroid enhancer along with GATA-1 and Ldb-1 in K562: evidence for the involvement of the SCL multi-protein complex in SCL regulation 177
5.13 Mapping interactions of proteins with HAT and HDAC activities across the SCL locus in K562 178
5.14 Mapping interactions of factor involved in the pre-initiation complex across the SCL locus in K562 179
5.15 Elf-1 binds to the SIL -1 region in K562, Jurkat and HPB-ALL 180
5.16 CTCF binding at the SCL locus in K562 and Jurkat 181
5.17 Discussion 182
  5.17.1 Patterns of histone acetylation across the human and mouse SCL locus 182
  5.17.2 Patterns of histone H3 methylation across the human and mouse SCL locus 183
  5.17.3 Histone H3 acetylation and methylation modifications can also be associated with inactive genes 185
  5.17.4 Nucleosome depletion occurs across the SCL locus in all human cell lines 186
  5.17.5 SCL may be self-regulated in a positive feedback loop 187
  5.17.6 Binding interactions of other transcription factors across the SCL locus 187
  5.17.7 Issues in interpreting ChIP-chip data 188
  5.17.8 Conclusions 190

Chapter 6 Analytical Approaches to Interpreting ChIP-chip Data at the SCL Locus

6.1 Introduction 191
6.2 Aims of this chapter 192
6.3 Sequence conservation at sites of histone H3 K4 dimethylation at the SCL locus in K562 and 416B cell lines 193
6.4 Sequence conservation at sites of histone H3 K4 trimethylation at the SCL locus in K562 and 416B cell lines 195
6.5 Relative levels of monomethylation to trimethylation of H3 lysine 4 correlate with transcriptional activity 197
6.6 Analysis of nucleosome density at the SCL locus Jurkat, HL60 and HPB-ALL 200
6.7 Visualization of ChIP-chip profiles for all histone modifications in SCL expressing and non-expressing cell lines 203
6.8 Hierarchical clustering of histone modifications distinguishes types of regulatory sequences and their activity in K562 205
6.9 Identification of a minimal set of histone modifications which delineates regulatory function at the SCL locus. 210
6.10 Discussion 217
  6.10.1 Sequence conservation at sites of di- and trimethylation of H3 K4 217
  6.10.2 Relative levels of mono- and trimethylation correlate with transcriptional activity 218
  6.10.3 Nucleosome depletion occurs across the SCL locus in all human cell lines 218
  6.10.4 Identification of a consensus histone code at the SCL locus 219

6.10.5 Conclusions 222

Chapter 7 Further Characterization of Regulatory Regions from Across the SCL Locus

7.1 Introduction 223
7.2 Aims of this chapter 224
7.3 Generating TreeView visualization profiles for the four cell lines 224
7.4 The SIL promoter and the SIL -10 region 225
7.5 The KCY promoter 228
7.6 The SCL promoters and their immediate 3’ flanking regions 229
7.7 Regulatory elements upstream of the SCL promoter 1a 232
    7.7.1 The -9/-10 region 233
    7.7.2 The -7 region: evidence for an active enhancer in K562 234
    7.7.3 The endothelial enhancer at the -3/-4 region 235
    7.7.4 The -12 region: a putative enhancer of SCL regulation 235
7.8 Regulatory regions downstream of the SCL gene 237
    7.8.1 The stem cell enhancer at the +20/+21 region 237
    7.8.2 The neural enhancer at the +23/+24 region 239
    7.8.3 The +39 region 239
    7.8.4 The MAP 17 promoter 241
    7.8.5 The erythroid enhancer at the +51/+53 region: evidence for sites of transcription downstream of the +51 region. 242
    7.8.6 The putative insulator element at the +57 region 245
7.9 Discussion 246
    7.9.1 Identification and characterization of putative novel regulatory elements 246
    7.9.2 Transcriptional activation of SCL in Jurkat 249
    7.9.3 MAP17 expression in K562 and Jurkat 249
    7.9.4 Conclusions 250

Chapter 8 Summary and Future work

8.1. Summary of the work presented in this thesis 251
    8.1.1. Construction and validation of the SCL genomic tiling path arrays (Chapter 3) ... 251
    8.1.2. Characterization of human haematopoietic cell lines (Chapter 3) 251
    8.1.3. Validation of the SCL arrays for ChIP-chip assays (Chapter 4) 252
    8.1.4. Applications and analysis of ChIP-chip assays across the SCL locus (Chapters 5, 6 & 7) 253
8.2. Future work 254
    8.2.1. Diagnostic studies 254
    8.2.2. Further profiling of DNA-protein interactions at the SCL locus 255
    8.2.3. Testing the histone code 256
    8.2.4. Further characterization of SCL regulatory regions: The +51 erythroid enhancer and +53 promoter 256

8.2.5. ChIP-on-chip studies: large scale analysis of mammalian genomes 258
8.2.6. Analysis of other biological features at the SCL locus 258
8.3. Final thoughts: SCL as a model of regulation for mammalian genomes 260

References 261

Appendices

Appendix 1 Sequences of Primer Pairs Used to Construct the Human SCL Tiling Path Array 285
Appendix 2 Sequences of Primer Pairs Used to Construct the Mouse SCL Tiling Path Array 297
Appendix 3A Sequences of Primer Pairs used to Amplify PCR Amplicons Containing GATA-1 Binding Sites (included as controls on the SCL tiling path array) 311
Appendix 3B Sequences of Primer Pairs used in the Real-Time PCR to Analyse SCL Expression in K562, Jurkat, HL60 and HPB-ALL 311
Appendix 4 Sequences of Primer Pairs used in the Real-Time PCR Verification of ChIP-chip data for Histone H3 K9/14 diacetylation 312
Appendix 5 Sequences of Primer Pairs used in the Real-Time PCR Verification of ChIP-chip data for GATA-1 314
Appendix 6 ChIP-chip profiles of histone H3 acetylation at specific lysine residues in K562, Jurkat, HL60 and HPB-ALL 315
Appendix 7 ChIP-chip profiles of histone H4 acetylation at specific lysine residues in K562, Jurkat, HL60 and HPB-ALL 321
Appendix 8 ChIP-chip profiles of histone H4 K5/8/12/16 tetraacetylation in K562, Jurkat, HL60 and HPB-ALL 326
Appendix 9 ChIP-chip profiles of histone H3 methylation at specific residues in K562, Jurkat, HL60 and HPB-ALL 327
Appendix 10 ChIP-chip profiles of histone H3 phosphorylation at serine 10 residue in K562, Jurkat, HL60 and HPB-ALL 332
Appendix 11 ChIP-chip profiles of histone H3 modifications at specific residues in 416B and mouse E14 ES cell lines 333
Appendix 12 Correlation graphs of non-coding sequence conservation with histone H3 acetylation and methylation 336
Appendix 13 Sequences of Primer Pairs to generate constructs for transfection assays 337
Appendix 14 Previous Publications

List of Figures

Figure 1.1: A schematic representation of the non-coding regulatory elements involved in the regulation of gene expression 4

Figure 1.2: Summary of the human general transcription factor interactions at a promoter 8

Figure 1.3: Schematic diagram of ChIP assay 24

Figure 1.4: Diagrammatic representation of the formation of a nucleosome 27

Figure 1.5: Histone modification map of mammalian histone N-terminal tails 30

Figure 1.6: Schematic of the human and mouse SCL loci 37

Figure 1.7: SCL mRNA expression during human haematopoietic differentiation. 39

Figure 1.8: A schematic representation of the known regulatory regions at the human and mouse SCL loci 50

Figure 3.1: Schematic diagram shows the approach adopted to make single-stranded PCR products 89

Figure 3.2: A schematic of the genomic regions across the human and mouse SCL loci. 91

Figure 3.3: Images of agarose gels after electrophoresis of second round PCR amplicons of SCL tiling path array 93

Figure 3.4: A histogram plot of a male versus female hybridisation on the 1st generation human SCL tiling path array 95

Figure 3.5: A histogram plot of a male versus female hybridisation on the final human SCL tiling path array 96

Figure 3.6: Analysis of SCL expression by real-time PCR 99

Figure 3.7: Array-CGH analysis of K562 cell line using the 1 Mb genomic array 100

Figure 3.8: Array-CGH analysis of Jurkat using the 1 Mb genomic array 101

Figure 3.9: Array-CGH analysis of HL60 using the 1 Mb genomic array 102

Figure 3.10: Array-CGH analysis of HPB-ALL using the 1 Mb genomic array 103

Figure 3.11: Array-CGH analysis of the four cell lines using the high resolution SCL genomic tiling path array 104

Figure 3.12: FISH images of metaphase spreads from the four human cell lines 107

Figure 4.1: ChIP-PCR based histone H3 acetylation profiles in K562 and 416B cell lines 115

Figure 4.2: The electrophoretic analysis of ChIP DNAs 117

Figure 4.3: A composite image of the human SCL genomic tiling array 118

Figure 4.4: Performance of the SCL array platform in ChIP-chip assays for H3 K9/14 diacetylation in K562 120

Figure 4.5: ChIP-chip profile of H3 K9/14 diacetylation across the human SCL locus in K562 121

Figure 4.6: SYBR green real-time PCR of ChIP material for H3 acetylation across the human SCL locus in K562 122

Figure 4.7: Comparison of histone H3 acetylation profiles across the SCL locus in human K562 and mouse 416B cell lines 124

Figure 4.8: Performance of the SCL array platform in ChIP-chip assays for GATA-1 in K562 128

Figure 4.9: ChIP-chip profile for GATA-1 across the human SCL locus in K562 129

Figure 4.10: SYBR green real-time PCR of ChIP material for GATA-1 across the human SCL locus 130
Figure 4.11: Conserved sequence analysis to identify conserved GATA and other transcription factors binding sites at SCL regulatory regions 131-133
Figure 4.12: Performance of the SCL array platform in ChIP-chip assays for histone H3 in K562 134
Figure 4.13: ChIP-chip profile for histone H3 across the human SCL locus in K562 135
Figure 4.14: Normalization of the array data for H3 K9/14 diacetylation with respect to underlying nucleosome occupancy in K562 137
Figure 5.1: The N-terminal tails of histones H3 and H4 144
Figure 5.2: Overall strategy employed to map regulatory interactions at the SCL locus using ChIP-chip 146
Figure 5.3: ChIP-chip profile for H3 K9/14 diacetylation across the SCL locus in K562 and Jurkat 151
Figure 5.4: ChIP-chip profiles for H3 K9/14 diacetylation across the SCL locus in HL60 and HPB-ALL 151
Figure 5.5: ChIP-chip profiles for histone H3 K9/14 diacetylation across SCL locus in 416B and mouse E14 ES cell lines 153
Figure 5.6: Composite ChIP-chip profiles for histone H3 acetylation at lysines 9, 14, 18 and 27 across the SCL locus in the four human cell lines 156
Figure 5.7: ChIP-chip profile of H2B tetra-acetylation at lysines 5/12/15/20 across the SCL locus in K562 158
Figure 5.8: ChIP-chip profiles for H4 K5/8/12/16 tetra-acetylation across the SCL locus in four human cell lines 159
Figure 5.9: ChIP-chip profiles for acetylation at lysines 5 and 16 of histone H4 across the SCL and MAP17 genes 161
Figure 5.10: ChIP-chip profiles for monomethylation of H3 K4 across the SCL locus in SCL expressing human and mouse cell lines 164
Figure 5.11: ChIP-chip profiles for monomethylation of H3 K4 across the SCL locus in SCL non-expressing human and mouse cell lines 166
Figure 5.12: ChIP-chip profiles of di- and trimethylated lysine 4 of H3 across the SCL locus in K562 and Jurkat 168
Figure 5.13: ChIP-chip profiles for di- and trimethylated lysine 4 of H3 across the SCL locus in HL60 and HPB-ALL 170
Figure 5.14: ChIP-chip profiles of di- and trimethylated lysine 4 of H3 across the SCL locus in 416B and mouse E14 ES cell lines 171
Figure 5.15: ChIP-chip profiles for histone H3 K79 trimethylation across the SCL locus in the four human cell lines 173
Figure 5.16: ChIP-chip profiles for H3 serine 10 phosphorylation across the 5' end of the SCL gene in four human cell lines 174
Figure 5.17: ChIP-chip profiles for histone H3 levels across the SCL locus in four human cell lines 176
Figure 5.18: Composite ChIP-chip profiles for GATA-1, SCL and Ldb-1 across the SCL locus in K562 177
Figure 5.19: Composite ChIP-chip profiles for p300, CBP and HDAC2 across the SCL locus in K562 178
Figure 5.20: ChIP-chip profile for Taf$_{II}$250 across the SCL locus in K562 179

Figure 5.21: Composite ChIP-chip profile for Elf-1 across the SCL locus in K562, Jurkat and HPB-ALL 180

Figure 5.22: ChIP-chip profile for CTCF binding across the SCL locus in K562 and Jurkat 182

Figure 6.1: Profiles showing trimethylation to monomethylation ratios in the human and mouse cell lines 198-199

Figure 6.2: A digital output of ChIP-chip profiles across the SCL locus for all the histone modifications in human cell lines 204

Figure 6.3: Hierarchical clustering defines active and inactive regions of chromatin across the SCL locus in K562 206

Figure 6.4: Identification and annotation of clusters of regulatory sequences across the SCL locus in K562 208-209

Figure 6.5: Identifying a minimal set of histone modifications for the SCL locus by comparing the “consensus” histone modifications of clusters 1 to 9 in K562 211

Figure 6.6: Hierarchical clustering of ChIP-chip data across the SCL locus in K562 using the minimal set of seven key histone modifications 214-215

Figure 6.7: A consensus histone code for the regulatory regions across the SCL locus 221

Figure 7.1: TreeView profiles at the 5’ end of the SIL gene and the SIL -10 region in the four human cell lines 226

Figure 7.2 Composite ChIP-chip profiles across the 5’ end of the SIL gene showing three distinct peaks 227

Figure 7.3: TreeView profiles at the 5’ end of the KCY gene in four human cell lines 229

Figure 7.4: TreeView profiles at the 5’ end of the SCL gene in four human cell lines............. 230

Figure 7.5: Composite ChIP-chip profiles across the 5’ end of the SCL gene........................ 231

Figure 7.6: TreeView profiles across the genomic regions upstream of the SCL promoter in the four cell lines 232

Figure 7.7: Composite ChIP-chip profiles across the SCL regulatory region located upstream of the SCL promoter 233

Figure 7.8: Transient luciferase reporter assays in K562 cells.. 236

Figure 7.9: Conserved transcription factor binding sites found at the SCL -12 region 237

Figure 7.10: TreeView profiles across the stem cell enhancer and the neural enhancer in the four human cell lines... 238

Figure 7.11: TreeView profiles and composite ChIP-chip profiles across the +39 region and MAP17 gene in the four human cell lines................... 240

Figure 7.12: TreeView profiles across the erythroid enhancer and the +57 region in four human cell lines 243

Figure 7.13: Composite ChIP-chip profiles for histone modifications at the +51/+53 regions in K562.. 244

List of Tables

Table 2.1: List of all the cell lines used for the study presented in this thesis 60

Table 2.2: Wash steps for hybridisations performed on the Tecan 76

Table 3.1: Summary of results of the construction of the SCL genomic tiling path array 91

Table 3.2: Cytogenetic characteristics and SCL expression patterns in the cell lines: K562, Jurkat, HL60 and HPB-ALL. 98

Table 4.1: The relationship between sequence conservation and histone H3 K9/14 diacetylation at the human and mouse SCL loci 127

Table 4.2: Sequence features of SCL array elements ranked according to their levels of histone H3 in K562 cells 136

Table 5.1: List of antibodies used in ChIP-chip assays 148-149

Table 5.2: Genomic regions across the SCL locus with significant ChIP enrichments for acetylation at lysines 9, 14, 18 and 27 of histone H3 in SCL expressing and non-expressing human cell lines 155

Table 5.3: Genomic regions across the SCL locus showing significant ChIP enrichments for histone H4 acetylation 160

Table 5.4: Genomic regions identified with mono-, di-, trimethylation of histone H3 at lysine 4 in human and mouse cell lines 167

Table 6.1: The relationship between sequence conservation and histone H3 K4 dimethylation at the SCL locus in K562 and 416B 195

Table 6.2: The relationship between sequence conservation and histone H3 K4 trimethylation at the SCL locus in K562 and 416B cell lines 196-197

Table 6.3: Sequence features of SCL array elements ranked according to their relative levels of histone H3 in the human cell lines 202

Table 7.1: A summary of the known regulatory regions and novel putative regulatory regions at the SCL locus in K562 246

Glossary of Abbreviations

ATP	adenosine triphosphate
bHLH	basic helix-loop-helix
bp	base pair
CBP	CREB binding protein
CGH	comparative genomic hybridisation
ChIP	chromatin immunoprecipitation
chip	microarray
C_ot	concentration x time (DNA reassociation kinetics)
CREB	cyclic AMP-responsive element-binding protein
CTD	C-terminal domain
CV	coefficient of variation
Cy3	cyanine-3
Cy5	cyanine-5
dATP	2' deoxyadenosine 5'-triphosphate
dCTP	2' deoxycytidine 5'-triphosphate
dGTP	2' deoxyguanosine 5'-triphosphate
DNA	deoxyribonucleic acid
dTTP	2' deoxythymidine 5'-triphosphate
ES cells	embryonic stem cells
ESTs	expressed sequence tags
FISH	fluorescence *in situ* hybridization
HAT	histone-acetyltransferase
HDAC	histone deacetylase
HMT	histone methyltransferase
HS	hypersensitive site
HSC	haematopoietic stem cell
IHGSC	International Human Genome Sequence Consortium
kb	kilobase (DNA)
MAP17	membrane associated protein
Mb	megabase (DNA)
mRNA	messenger ribonucleic acid

nc-RNA	non coding RNA
ORF	open reading frame
PCR	polymerase chain reaction
PIC	pre-initiation complex
RNA	ribonucleic acid
RNA Pol II	ribonucleic acid polymerase II
RT-PCR	real-time polymerase chain reaction
SCL	stem cell leukaemia
SD	standard deviation
SIL	SCL interrupting locus
TAF	TFIID associating factor
T-ALL	T-cell acute lymphoblastic leukaemia
TBP	TATA-binding protein
TCR	T-cell receptor
TF	transcription factor
tRNA	transfer RNA
UTR	untranslated region