

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

**Does DNA topography coordinate intra-
and inter-chromosomal Galactose (*GAL*)
gene expression?**

**This thesis is presented in partial fulfillment of the
requirements for the degree of Masters of Science (MSc) in
Molecular Biology at Massey University, Albany, New
Zealand**

Shuqiao Yu

2008

Abstract

For a long time, DNA had been considered as a stabilized, rigid, and “linear” structure, which acts as a platform for molecular regulators to function. However, genome structure in living cells is far more complex than the linear representation of the primary DNA sequence implies. This thesis aims to investigate whether the position of a gene within the genome plays a role in the regulation of its activity. The galactose (*GAL*) gene family of *Saccharomyces Cerevisiae* is used as a model. This gene family enables yeast cells to utilize galactose as an alternative carbon source; and it consists of structural and regulatory genes. Structural genes *GAL1*, *GAL10* and *GAL7* exist in a cluster on yeast chromosome II. The products of the regulatory genes, *GAL3*, *GAL4*, and *GAL80*, regulate the expression of the *GAL* structural genes, depending on the availability of carbon sources. Specifically, *GAL* gene expression is repressed by glucose, paused for induction (noninduced) by glycerol/lactate, and fully induced by galactose.

The aim of this project was to study the relative position of the *GAL* structural genes within the nucleus, and whether any chromosomal interactions at the *GAL* locus help to regulate their activation. These were tested in accordance with the expression status of the *GAL* genes (*i.e.* repressed, noninduced or induced). Followed confirmation of the existence of any chromosomal interactions, protein/protein complexes that mediate these interactions were attempted to identify.

The methods applied in this project were Chromosome Conformation Capture (3C) and Circular Chromosome Conformation Capture (4C), which applied in combination to map the positions of the *GAL* genes in the context of the overall genome structure. The results indicated that the *GAL* locus on chromosome II was divided into two “interaction zones”. DNA loops formed around these interaction zones to form an S-shape structure in a carbon source-independent manner. Two novel inter-chromosomal interactions between chromosomes II and XVI, *i.e.* *SVL3-GAL7* and *HOS1-GAL10*, were also identified. Although these interactions occurred regardless of the *GAL* gene activities, it was suggested by real-time PCR that the interaction frequency for *SVL3-GAL7* declined as the *GAL* genes being activated. Unfortunately no protein/protein complexes were identified to play an important role in mediating either intra- or inter-chromosomal interactions.

Future work will be needed to identify the protein/protein complexes that play a role in mediating the S-Shape structure at the *GAL* locus and the two inter-chromosomal interactions. Additional works could also focus on the understanding of the functional implication of the interactions between chromosomes II and XVI.

Acknowledgement

Firstly and most importantly I would like to say a big thank you to my supervisor, Professor David Lambert, who helped me through the most difficult time in my Masters studies. Without him, this thesis would not exist. There is no way I could fully express my appreciation. I guess the finalization and completion of this thesis is the best return for his unconditional supports.

I would also like to take this opportunity to thank you a lot of people who have been offering me enormous helps and supports throughout the years. Without you, this thesis would not exist either. I would like to give my sincere appreciation to Tim Coolbear and Alan Clark for their generous supports, especially Tim, thanks for being there for me at all time. Justin O'Sullivan for his helps in early project development and methodology and financial supports. Leon Huynan for offering me with great suggestions and feedback for both my project and this thesis. A lot of other people in the lab, Jarod, Eli, Martina, Dharani, Katie, Tom, etc., thank you so much for the happy time we shared together. Also Rebecca, Tracey and Daniela who have already left. My friends Sha, Wei, Soyeon, Sara, Kylie, Nancy and a lot more if I missed your names. Thank you so much for the continuous supports.

Finally, I would like to thank my parents for believing in me from the very beginning and for their unconditional love and support, both financially and emotionally. My final and special thank is to Kai, who has been listening to me, comforting me when I was down and most importantly, share the best and worst moments with me. A very deep thank you!

Lists of Figures

Figure 1: Analogue model for a transcription factory.....	5
Figure 2: Analogue models for DNA loop formation of the murine α -globin locus.....	9
Figure 3: The galactose utilization pathway.....	11
Figure 4: <i>GAL</i> gene regulation system.....	12
Figure 5: Schematic representation of the 3C and 4C assays.....	21
Figure 6: Schematic representation of the 4C pulldown assay.....	28
Figure 7: Diagrammatical representation of primers and probe for <i>Taqman</i> [®] real-time PCR.....	31
Figure 8: Diagrammatic representation of the <i>Taqman</i> [®] Real-time PCR Application.....	32
Figure 9: Diagrammatical representation of <i>GAL</i> structural genes 1, 7 and 10 on chromosome II.....	35
Figure 10: Diagrammatical representation of primer design at the <i>GAL</i> Locus on chromosome II.....	37
Figure 11: Positive control testing.....	37
Figure 12: Preliminary 3C analyses of interactions at the <i>GAL</i> locus.....	39
Figure 13: Diagrammatical representation of primer design at the <i>GAL</i> locus on chromosome II for 4C analyses.....	40
Figure 14: 4C nested PCR analyses of interactions at the <i>GAL</i> locus of chromosome II.....	41
Figure 15: PCR analyses for direct sequencing.....	43
Figure 16: Diagrammatical representation of the direct sequencing Results.....	43
Figure 17: Diagrammatical representation of the clonal sequencing Results.....	46
Figure 18: Diagrammatical representation of the “walking-down” strategy in 3C studies at the <i>GAL</i> locus.....	48
Figure 19: Diagrammatical representation of the “two interaction zone” at the <i>GAL</i> locus.....	50
Figure 20: Diagrammatical representation of the “long distance” 3C Analyses at the <i>GAL</i> locus.....	51
Figure 21: Modelling of an S-shape structure at the <i>GAL</i> locus of chromosome II.....	53
Figure 22: Diagrammatical representation of the critical interactions at S-shape structure.....	54
Figure 23: Gal80p deletion confirmation.....	55
Figure 24: 3C analyses on Gal80p for intra-chromosomal interactions.....	56
Figure 25: 3C analyses on Rat1p for intra-chromosomal interactions.....	58
Figure 26: Inter-chromosomal interactions confirmation.....	62
Figure 27: 3C PCR analysis confirming the isolation of Gal4p in 4C	

pull-down experiment.....	64
Figure 28: 3C analyses on Gal4p for inter-chromosomal interactions.....	65
Figure 29: 3C analyses on Gal80p for inter-chromosomal interactions.....	66
Figure 30: 3C analyses on Rat1p for inter-chromosomal interactions.....	67
Figure 31: Diagrammatical representation of internal primers design at the <i>GAL10</i> gene.....	69
Figure 32: <i>SyBr</i> green real-time PCR analyzed DNA template concentration into each PCR reaction.....	71
Figure 33: Column graph illustrates the interaction frequency for <i>SVL3-GAL7</i> under various induction conditions.....	72
Figure 34: Location of the DNase I-hypersensitive sites between <i>GAL10</i> and <i>GAL1</i> on chromosome II.....	75
Figure 35: Combinational model of chromosomes II and XVI.....	77

Lists of Tables

Table 1: <i>S. cerevisiae</i> strains and the plasmid used in this study.....	17
Table 2: Information of the <i>GAL</i> structural genes on chromosome II.....	33
Table 3: Summary of the cloning sequencing at the <i>GAL</i> locus of chromosome II.....	44
Table 4: Calculation of <i>SVL3-GAL7</i> interaction frequency.....	73

Abbreviations

In addition to the chemical symbols shown in the Periodic Table of Elements and the *Systeme international d'unités* (SI), the following abbreviations are used:

SC	Synthetic complete
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
DMSO	Dimethyl sulfoxide
EDTA	Ethylenediaminetetraacetic acid
PCR	Polymerase chain reaction
PBS	Phosphate buffered saline
PMSF	Phenylmethanesulfonyl fluoride
SDS	Sodium dodecyl sulfate
SOC	Super optimal catobolite
BSA	Bovine serum albumin
TE	Tris EDTA
TBE	Tris borate EDTA.
TEMED	Tetramethylethylenediamine
LPA	Linear acrylamide
H	Histidine
Kb	Kilobase
w/v	Weight/volume
v/v	Volume/volume
RT	Room temperature
O/N	Overnight
Min	Minute
S	Second
U	Unit
3C	Chromosome Conformation Capture
4C	Circular Chromosome Conformation Capture
CHO	Carbohydrate
C	Cut/restriction enzyme digested sample
CL	Cut/restriction enzyme digested and ligated sample

Table of Contents

Chapter 1----Introduction.....	1
1.1 Preview.....	1
1.2 Nuclear Structure, Chromosome Territories and Gene Positioning.....	2
1.2.1 Gene Organization within Chromosomes and Regulation of Transcription..	3
1.2.2 Inter-Chromosomal Association and Regulation of Transcription	6
1.3 Gene Looping and Transcription Regulation	7
1.4 Galactose Gene (<i>GAL</i>) Family in <i>Saccharomyces Cerevisiae</i>	10
1.5 Methodology	15
1.6 Research Aims and Objectives	16
Chapter 2----Materials and Methods.....	17
2.1 Cultures and Plasmid	17
2.1.1 Cultures and Plasmid used.....	17
2.1.2 Growth of Fungal Cultures	18
2.2 Nucleic Acid Extractions	18
2.2.1 Genomic DNA.....	18
2.2.2 Plasmid Preparation.....	19
2.3 Molecular Techniques	19
2.3.1 Polymerase Chain Reaction (PCR)	19
2.3.2 Primer Design and Restriction Enzyme Identification.....	20
2.3.3 Gel Electrophoresis.....	20
2.3.4 DNA Extraction from Agarose Gels	20
2.3.5 DNA Purification.....	20
2.3.6 Chromosome Conformation Capture (3C) Analysis	21
2.3.6.1 Chromatin Extraction.....	22
2.3.6.2 Chromatin Digestion.....	22
2.3.6.3 Chromatin Ligation.....	23
2.3.6.4 Chromatin Reverse Cross-linking.....	23
2.3.6.5 DNA Isolation and Purification.....	23
2.3.6.6 3C PCR Analysis.....	23
2.3.6.7 3C PCR Positive Controls.....	24
2.3.7 Circular Chromosome Conformation Capture (4C) Analysis.....	24
2.3.7.1 4C PCR Analysis.....	25
2.3.7.2 4C Generation of Library Inserts	25
2.3.7.3 Cloning-Digestion and Dephosphorylation of pUC18	25
2.3.7.4 Ligation of 4C Library inserts to Dephosphorylated pUC18.....	26
2.3.7.5 Transformation of Electro-competent <i>E. coli</i> DH5a Cells.....	26
2.3.7.6 Analysis of Cloned 4C Library	27
2.3.8 4C Pulldown.....	27
2.3.8.1 Protein Isolation (“Pulldown”).....	28
2.3.8.2 PCR Analyses for 4C Pulldown.....	29
2.4 Determination of Interaction Frequency by Real-time PCR	30
2.4.1 SyBr Green Real-time PCR Application.....	30
2.4.2 Taqman [®] Real-time PCR.....	30
2.5 Bioinformatics.....	33

2.5.1 BLAST Searches	33
2.5.2 Sequence Alignments	33

Chapter 3----Intra-chromosomal Interactions at the <i>GAL</i> Locus of Chromosome II.....	34
3.1 Introduction	34
3.2 Chromatin Isolation and Preliminary 3C Studies	35
3.2.1 Restriction Enzyme Selection	35
3.2.2 Primer Design and Positive Control for 3C	36
3.2.3 Preliminary 3C Analyses	38
3.3 Switch to 4C Analysis	39
3.3.1 4C Principle and Application.....	39
3.3.2 4C Results at the <i>GAL</i> Locus	40
3.3.3 Identify of the Unknown Sequence Derived from 4C Analyses.....	42
3.4 3C Studies of Intra-chromosomal Interactions	47
3.4.1 3C Analyses on Adjacent Fragments.....	47
3.4.2 3C Analyses on Non-adjacent Fragments.....	50
3.5 Exploration of the Formation of S-shape Structure	53
3.5.1 Experiment 1-- $\Delta gal80p$ Strain	54
3.5.2 Experiment 2-- <i>rat1-1</i> Temperature-sensitive Strain.....	57
3.6 Summary	59

Chapter 4----Inter-chromosomal Interactions at the <i>GAL</i> Locus of Chromosome II.....	60
4.1 Introduction	60
4.2 Interactions Between the <i>GAL</i> Locus and Other Chromosome	60
4.2.1 The Promoter of <i>GAL7</i> Interacts with the <i>SVL3</i> Gene	60
4.2.2 The Promoter of <i>GAL1</i> Interacts with the <i>HOS1</i> Gene	61
4.2.3 3C Confirmation of <i>GAL/SVL3/HOS1</i> Interactions	61
4.3 Exploration of the Mediation of Interactions	63
4.3.1 Experiment 1--TAP-tagged Gal4p	63
4.3.2 Experiment 2-- $\Delta gal80p$ Strain	65
4.3.3 Experiment 3-- <i>rat1-1</i> Temperature-sensitive Strain.....	66
4.4 Quantitative Real-time PCR Analysis	68
4.4.1 <i>SyBr</i> Green Real-time PCR.....	68
4.4.2 <i>Taqman</i> [®] Real-time PCR.....	70
4.5 Summary	73

Chapter 5----Discussion and Summary.....	74
5.1 Studies on DNA Structural Changes at the <i>GAL</i> Locus.....	74
5.1.1 Non-linear DNA Structure at <i>GAL</i> Locus on Chromosome II.....	74
5.1.2 Interactions Between Chromosomes II and XVI	77
5.2 Functional Implications of Changes in <i>GAL</i> Gene Positioning	79
5.2.1 Exploration of S-shape Structure Formation at the <i>GAL</i> Locus.....	79
5.2.2 Exploration of the Inter-chromosomal Interactions at the <i>GAL</i> Locus	80
5.3 Real-time PCR Studies on Interaction Frequency.....	81
5.3.1 Interaction Frequency for the <i>SVL3-GAL7</i> Interaction.....	81

5.4 Methodology	82
5.5 Significance	83
5.6 Summary	84
5.7 Future studies	84
Appendices:	86
Appendix 1.0 Media Preparation	86
Appendix 2.0 Common Buffers and Solutions.....	87
Appendix 2.1 Buffers Used in 3C, 4C and 4C Pulldown.....	87
Appendix 3.0 Primers.....	89
Appendix 4.0 pUC18 Vector	90
Appendix 5.0 Preparation of Bacterial Cells for Electroporation	91
Appendix 6.0 Amplification of the ~2kb Regions for the Production of Positive Controls in 3C Analysis	92
Appendix 7.0 Primer testing in 4C Analysis.....	93
Appendix 8.0 Cross-Linking Time Determination	94
Appendix 9.0 BLAST Results	96
Appendix 9.1 BLAST Analysis Showed a Matching Sequence to the <i>SVL3</i> ORF	96
Appendix 9.2 BLAST Analysis Showed a Matching Sequence to the <i>HOS1</i> ORF	96
Appendix 9.3 BLAST Analysis Showed a Matching Sequence to the <i>SVL3</i> ORF and Sequence of Chromosome II	97
Appendix 10.0 Real-time PCR Standards and Data	98
Appendix 10.1 <i>SyBr</i> Green and <i>Taqman</i> [®] Standards	98
Appendix 10.2 <i>SyBr</i> Green PCR Data.....	98
Appendix 10.3 <i>Taqman</i> [®] PCR Dissociation Curve.....	99
Appendix 10.4 <i>Taqman</i> [®] PCR Data	100
Appendix 10.5 <i>Taqman</i> [®] PCR Data Calculation	101
Appendix 11.0 GMO Approval	102