



Murdoch
UNIVERSITY

The Role of Genomic Copy Number Variation (CNV) in Osteoporosis

Kate Connolly

Bachelor of Forensics and Bachelor of Science (Triple Major)
School of Biological Sciences and Biotechnology
Murdoch University
2012

Principle Supervisor: A/Professor Scott Wilson

Department of Endocrinology and Diabetes
Sir Charles Gairdner Hospital

Co-Supervisor: Professor Roger Price

Department of Endocrinology and Diabetes
Sir Charles Gairdner Hospital

Responsible Academic: A/Professor Robert Mead

School of Biological Sciences and Biotechnology
Murdoch University

This thesis was submitted in partial fulfilment of the requirements for the award of Honours Degree in Molecular Biology at Murdoch University, Western Australia

Table of Contents

TABLE OF CONTENTS.....	I
DECLARATION.....	IV
ACKNOWLEDGEMENTS.....	V
LIST OF ABBREVIATIONS.....	VI
INDEX OF FIGURES.....	IX
INDEX OF TABLES.....	X
ABSTRACT.....	XII
CHAPTER 1: INTRODUCTION.....	1
1.1 THE IMPACT OF OSTEOPOROSIS.....	2
1.2 THE OSTEOPOROTIC DISEASE PROCESS.....	5
1.2.1 BONE BIOLOGY.....	5
1.2.1.1 Bone Remodelling.....	5
1.2.1.2 Bone Regulation.....	6
1.2.2 PRIMARY RISK FACTORS FOR OSTEOPOROSIS.....	8
1.3 EARLY INTERVENTION OF OSTEOPOROSIS.....	10
1.3.1 DETECTION OF FRACTURE RISK VIA BMD.....	10
1.3.2 PREVENTION OF OSTEOPOROSIS.....	11
1.4 THE GENETIC CONTRIBUTION TO OSTEOPOROSIS.....	13
1.4.1 COMMON COMPLEX PATTERN OF INHERITANCE.....	13
1.4.2 IDENTIFICATION OF GENETIC FACTORS THAT ARE ASSOCIATED WITH DISEASE.....	14
1.4.3 EVIDENCE SUPPORTING THE HERITABILITY OF OSTEOPOROSIS.....	16
1.5 HUMAN GENETIC VARIATION.....	18
1.5.1 FUNCTIONAL CONSEQUENCE OF HUMAN GENETIC VARIATION.....	19
1.5.1.1 Transcriptional Control of Gene Expression.....	19
1.5.2 CHARACTERISATION OF HUMAN GENETIC VARIATION.....	21
1.5.2.1 Single Nucleotide Polymorphisms (SNPs).....	22
1.5.2.2 Structural Variation.....	23
1.6 COPY NUMBER VARIATION (CNV).....	25
1.6.1 MECHANISM OF CNV FORMATION.....	25
1.6.2 FUNCTIONAL CONSEQUENCE OF CNV.....	26
1.6.3 CNVs AND DISEASE.....	27
1.7 STATEMENT OF THESIS.....	30
1.7.1 GENERAL PURPOSE OF RESEARCH.....	30
1.7.2 HYPOTHESIS.....	30
1.7.3 AIMS OF STUDY.....	30
CHAPTER 2: MATERIALS AND METHODS.....	32
2.1 MATERIALS.....	33

2.1.1 SELECTION OF CNVs FOR STUDY	33
2.1.2 gDNA PREPARATION	33
2.1.3 qPCR (REAL-TIME QUANTITATIVE POLYMERASE CHAIN REACTION)	34
2.1.4 ANALYSIS OF EXPERIMENTAL DATA	35
2.1.5 GENERAL LABORATORY PROTOCOL	35
2.2 METHODS	36
2.2.1 SELECTION OF CNVs FOR STUDY	36
2.2.1.1 <i>Identification of Potential Osteoporosis Genes</i>	36
2.2.1.2 <i>Identification of CNVs</i>	37
2.2.1.3 <i>Critical Evaluation of CNVs</i>	38
2.2.2 SAMPLE PREPARATION	39
2.2.2.1 <i>Study Subjects</i>	39
2.2.2.2 <i>gDNA Dilution</i>	40
2.2.3 DETECTION OF CNV	41
2.2.3.1 <i>Overview of qPCR</i>	42
2.2.3.2 <i>Overview of TaqMan Assay Detection Chemistry</i>	43
2.2.3.3 <i>Selection of TaqMan Assays used in Study</i>	44
2.2.3.4 <i>CNV Genotyping via qPCR</i>	47
2.2.4 QUANTITATION OF CNVs	48
2.2.4.1 <i>Overview of CopyCaller</i>	48
2.2.5 CRITICAL EVALUATION OF EXPERIMENTAL DATA	51
2.2.5.1 <i>Tests of Reliability and Reproducibility</i>	51
2.2.5.2 <i>Prioritisation of Samples with 'Unconfirmed' CN from First Round of Genotyping</i>	52
2.2.5.3 <i>Prioritisation of Samples with 'Unconfirmed' CN from Second Round of Genotyping</i>	53
2.2.6 STATISTICAL ANALYSIS OF EXPERIMENTAL DATA	54
2.2.6.1 <i>Consolidation of Genotype and Phenotype Data</i>	54
2.2.6.2 <i>Tests of Normality via SPSS</i>	55
2.2.6.3 <i>Tests of Association via SPSS</i>	56
CHAPTER 3: ANALYSIS OF ESR1 CNV AND BMD	59
3.1 INTRODUCTION	60
3.2 MATERIALS AND METHODS	66
3.3 RESULTS	67
3.3.1 QUALITY CONTROL	67
3.3.2 CNV GENOTYPE FREQUENCY DISTRIBUTIONS	68
3.3.3 SPSS TESTS OF NORMALITY	69
3.3.3.1 <i>ESR1 CNV and Raw BMD Measurements (g/cm²)</i>	69
3.3.3.2 <i>ESR1 CNV and Age-Adjusted BMD Z-Scores</i>	73
3.3.4 ASSOCIATION ANALYSIS	76
3.3.4.1 <i>ESR1 CNV and Raw BMD Measurements (g/cm²)</i>	76
3.3.4.2 <i>ESR1 CNV and Age-Adjusted BMD Z-Scores</i>	77
3.3.4.3 <i>Effect of Covariates on ESR1 CNV and BMD Measurements (g/cm²)</i>	78
3.4 DISCUSSION	81
CHAPTER 4: ANALYSIS OF AXIN1 CNV AND BMD	84
4.1 INTRODUCTION	85
4.2 MATERIALS AND METHODS	90

4.3 RESULTS	91
4.3.1 QUALITY CONTROL	91
4.3.2 CNV GENOTYPE FREQUENCY DISTRIBUTIONS.....	93
4.3.3 SPSS TESTS OF NORMALITY	93
4.3.3.1 AXIN1 and Raw BMD Measurements (g/cm ²).....	93
4.3.3.2 Age-Adjusted BMD Z-Scores.....	97
4.3.4 ASSOCIATION ANALYSIS	100
4.3.4.1 AXIN1 CNV and Raw BMD Measurements (g/cm ²).....	100
4.3.4.2 AXIN1 CNV and Age-Adjusted BMD Z-Scores.....	101
4.3.4.3 Effect of Covariates on AXIN1 CNV and BMD Measurements (g/cm ²).....	102
4.4 DISCUSSION	105
CHAPTER 5: GENERAL DISCUSSION AND CONCLUSION	108
5.1 GENERAL DISCUSSION.....	109
5.1.1 CNV DETECTION AND QUANTITATION.....	110
5.1.2 CNV ASSOCIATION WITH BMD	113
5.1.3 RECOMMENDATIONS FOR FURTHER STUDY.....	115
5.2 CONCLUSION.....	117
REFERENCES.....	119
APPENDIX	129

Declaration

This thesis is my own original work and does not incorporate the material of others without proper acknowledgement and appropriate referencing.

Signed: _____

Date: _____

Kate Connolly, 2012

Word Count: 30,931

Acknowledgements

There are a number of people to whom I would like to extend my sincere gratitude and thanks. I would not have been able to complete this project without the assistance and support of these people.

Firstly, I would like to thank my supervisor, A/Prof Scott Wilson, for providing me the opportunity to undertake this project. Thank you for your insight, guidance, and especially for your ability to motivate when I needed it.

I would like to extend a special thank you to A/Prof Bob Mead, for his ongoing guidance and good advice throughout my years at Murdoch. Thank you for your encouragement and support.

I would like to thank Shelby Chew, for her patience, assistance, and friendship throughout this project. Thank you for assisting with my constant enquiries and difficulties, and for being a gracious mentor.

To Suzanne Brown, thank you for assisting with my statistical analyses. Your expertise was greatly appreciated.

To my fellow Honours students and to all the staff at the SCGH Endocrinology Department, thank you for making the lab a pleasant and inspiring place to work and learn.

To Richard Allcock and to the staff at the Lotterywest State Biomedical Facility Genomics at RPH, thank you for the use of the ABI 7900HT thermal cycler.

And finally, a heartfelt thank you to my family and friends for their love and support throughout my studies and my time working on this project.

List of Abbreviations

Abbreviation	Definition
α	Alpha
B	Beta
κ	Kappa
n	Nano
μ	Micro
ABI	Applied Biosystems by Life Technologies
ABS	Australian Bureau of Statistics
aCGH	array-Comparative Genomic Hybridisation
ALP	Alkaline Phosphatase
ANOVA	Analysis of Variance
ANCOVA	Analysis of Covariance
APC	Adenomatous Polyposis Coli
AXIN1	Axin-1 gene
BAC	Bacterial Artificial Chromosome
BMD	Bone Marrow Density
Bp	Base pairs
CN	Copy Number
CNV	Copy Number Variant
C_T	Cycle Threshold
COLIA1	Collagen Type I Alpha 1 gene
DGV	Database of Genomic Variants
df	Degrees of Freedom
Dvl	Dishevelled
DXA	Dual-energy X-ray Absorptiometry
EDTA	Ethylenediaminetetra-acetic Acid
ESR1	Oestrogen receptor 1 gene

FBAT	Family-Based Association Tests
FRET	Förster Resonance Energy Transfer
Fz	Frizzled
H1RNA	RNase P RNA component H1
gDNA	Genomic DNA
GSK-3	Glycogen Synthase Kinase-3
GWAS	Genome-Wide Association Study
HCl	Hydrochloric Acid
HGA	Human Genome Assembly
IF	Interleukin
LRP5	Low-density Lipoprotein Receptor-related Protein 5
MAF	Minor Allele Frequency
MGB	Minor Groove Binder
MSC	Mesenchymal Stem Cells
NaOH	Sodium Hydroxide
NHS	National Health Survey
NTC	Non-Template Control
OPG	Osteoprotegerin
ORF	Open Reading Frame
PTH	Parathyroid Hormone
qPCR	Quantitative Real-Time Polymerase Chain Reaction
Q-Q	Quantile-Quantile
RAB11FIP3	RAB11 family interacting protein 3 (class II) gene
RANK	Receptor Activator of Nuclear factor κ B
RANKL	Receptor Activator of Nuclear factor κ B Ligand
RFLP	Restriction Fragment Length Polymorphism
RPPH1	RNase P RNA component H1 gene
RUNX2	Runt-related gene 2

SCGH	Sir Charles Gairdner Hospital
SD	Standard Deviation
SNP	Single Nucleotide Polymorphism
SOST	Sclerostin
SPSS	Statistical Package for the Social Sciences
TF	Transcription Factor
TGF- β	Transforming Growth Factor-Beta
TNF- α	Tumour Necrosis Factor-Alpha
Tris	Trizma Base
VDR	Vitamin D Receptor gene
VNTR	Variable Number Tandem Repeat
UCSC	University of California Santa Cruz

Index of Figures

Figure Title	Page no.
Figure 1: Routine Calculations for gDNA Dilution Volumes	41
Figure 2: Summary of TaqMan CN Assays Selected for Study	45
Figure 3: Genomic Location of ESR1 CNV Variant_4512 in Relation to ESR1 and C6orf97	64
Figure 4: Selected Examples of Normal Q-Q Plots for ESR1 CNV and BMD (g/cm ²)	71
Figure 5: Selected Examples of Box Plots for ESR1 CNV and BMD (g/cm ²)	72
Figure 6: Selected Examples of Normal Q-Q Plots for ESR1 CNV and BMDZ	74
Figure 7: Selected Examples of Box Plots for ESR1 CNV and BMDZ	75
Figure 8: Genomic Location of AXIN1 CNV Variant_4912 in Relation to AXIN1 and RAB11FIP3	88
Figure 9: Selected Examples of Normal Q-Q Plots for AXIN1 CNV and BMD (g/cm ²)	95
Figure 10: Selected Examples of Box Plots for AXIN1 CNV and BMD (g/cm ²)	96
Figure 11: Selected Examples of Normal Q-Q Plots for AXIN1 CNV and BMDZ	98
Figure 12: Selected Examples of Box Plots for AXIN1 CNV and BMDZ	99

Index of Tables

Table Title	Page no.
Table 1: Databases of Genomics and Bioinformatics Used to Select CNVs for Study	33
Table 2: Source of gDNA Samples for Study Subjects	33
Table 3: Reagents for TE (Tris-EDTA) Buffer	33
Table 4: TaqMan CN Assays Used to Amplify and Detect CNVs	34
Table 5: qPCR Reagents for TaqMan CN Assays	34
Table 6: qPCR Instruments and Equipment	34
Table 7: Method used for CN Prediction	35
Table 8: Statistics Packaged Used to Test Association of CNV and BMD	35
Table 9: General Equipment Used in Laboratory	35
Table 10: Short-List of Potential Candidate CNVs for Study	38
Table 11: Clinical BMD Phenotype Data for Cohort Study	40
Table 12: qPCR Reaction Mixture for TaqMan CN Assays	48
Table 13: qPCR Thermal Cycling Conditions for TaqMan CN Assays	48
Table 14: Summary of Confirmed ESR1 CNV Genotypes	69
Table 15: Shapiro-Wilk Test of Normality for ESR1 CNV and BMD (g/cm^2)	70
Table 16: Shapiro-Wilk Test of Normality for ESR1 CNV and BMDZ	73
Table 17: Levene's Test for Homogeneity of Variances for ESR1 CNV and BMD (g/cm^2)	76
Table 18: ANOVA Test of Association for ESR1 CNV and BMD (g/cm^2)	77
Table 19: Levene's Test for Homogeneity of Variances for ESR1 CNV and BMDZ	77
Table 20: ANOVA Test of Association for ESR1 CNV and BMDZ	78
Table 21: Levene's Test for Homogeneity of Variances of ESR1 CNV and BMD (g/cm^2), Controlled for Covariates	79
Table 22: ANCOVA Test of Association for ESR1 CNV and BMD (g/cm^2), Controlled for Covariates	80
Table 23: Summary of Confirmed AXIN1 CNV Genotypes	93
Table 24: Shapiro-Wilk Test of Normality for AXIN1 CNV and BMD (g/cm^2)	94

Table 25: Shapiro-Wilk Test of Normality for AXIN1 CNV and BMDZ	94
Table 26: Selected Examples of Normal Q-Q Plots for AXIN1 CNV and BMDZ	95
Table 27: Selected Examples of Box Plots for AXIN1 CNV and BMDZ	96
Table 28: Levene's Test for Homogeneity of Variances for AXIN1 CNV and BMD (g/cm ²)	100
Table 29: Independent-Sample T-Test of Association for AXIN1 CNV and BMD (g/cm ²)	101
Table 30: Levene's Test for Homogeneity of Variances for AXIN1 CNV and BMDZ	101
Table 31: Independent-Sample T-Test of Association for AXIN1 CNV and BMDZ	102
Table 32: Levene's Test for Homogeneity of Variances for AXIN1 CNV and BMD (g/cm ²), Controlled for Covariates	103
Table 33: ANCOVA Test of Association for AXIN1 CNV and BMD (g/cm ²), Controlled for Covariates	104

Abstract

Copy number variation (CNV) is a relatively novel source of genetic variation, involving the duplication or deletion of segments of genomic DNA (gDNA) sequence, thereby changing the original number of DNA copies. It is currently gaining widespread recognition from the scientific community, and it is anticipated to play a major role in the aetiology of human diseases. However, the extent of its contribution to phenotypic diversity, in terms of individual susceptibility to disease, remains to be elucidated. Nonetheless, recent studies have indicated that common complex disease phenotypes, such as osteoporosis, might be highly susceptible to CNV.

Osteoporosis is a common and debilitating skeletal condition, imposing significant clinical and socioeconomic consequences. The disease is characterised by fragile bones that are susceptible to fracture due to deregulated bone remodelling, where bone loss exceeds bone formation. Being a common complex disease, osteoporosis risk is largely determined by the effect of environmental factors on genetic variants. Moreover, identification of the genetic variants associated with osteoporosis is widely anticipated for the contribution it will make towards the development of improved measures of disease intervention.

Recent genome-wide association studies (GWASs) have identified that the genes oestrogen receptor 1 (ESR1) and Axin 1 (AXIN1) potentially play major roles in bone regulation. In addition, evidence highlights their involvement in key biological processes that regulate bone turnover. Specifically, ESR1 mediates the response of bone marrow-derived cells to oestrogen and it has been demonstrated that oestrogen inhibits bone loss, while AXIN1 inhibits *Wnt* signal transduction and it has been demonstrated that *Wnt* proteins promote bone growth. Furthermore, several large-scale analysis projects firmly implicate genetic

variations of both genes with bone marrow density (BMD), which is the surrogate phenotype of osteoporosis. Therefore, ESR1 and AXIN1 are both recognised candidates for the genetic regulation of osteoporosis risk.

This study investigated the potential effect of two novel CNVs of the genes ESR1 and AXIN1, Variant_4512 and Variant_4912, respectively, in relation to BMD in a population cohort study of Caucasian women, between the ages of 18 and 83, from Australia and the UK. Subjects were genotyped for both CNVs, respectively, using real-time quantitative PCR (qPCR) combined with TaqMan chemistry, and the copy number (CN) quantitation software, CopyCaller. Subjects were then examined for evidence of association between both CNVs and three different BMD phenotypes, 1) raw measurement (g/cm^2), 2) age-adjusted Z-score, and 3) controlled for several covariates, at three common skeletal locations of osteoporotic fracture, 1) lumbar spine, 2) total hip, and 3) femoral neck.

This study confirmed the presence of ESR1 CNV and AXIN1 CNV in the analysed subject cohort, as indicated by the observation of three distinct CNV genotypes for each, representing CN loss (CN1) and CN gain (CN3) from the expected wild-type CN in the human diploid genome (CN2). This study found no evidence of association between both CNVs and BMD ($p = > 0.05$) in the analysed subject cohort. Therefore, the hypothesis tested in this study, that CNV is associated with BMD, was not supported. As a result, it would appear that the ESR1 CNV Variant_4512 and the AXIN1 CNV Variant_4912 are unlikely to play a major role in the pathogenesis of osteoporosis in Caucasian women. However, replication studies and further research would be required to accurately validate this, since this study was subject to numerous limitations which may have influenced the findings, such as low statistical power, technical difficulties, limiting experimental reagents, and time constraints.

In addition, there is evidence from previous studies implicating intron 1 and the 5' end of ESR1 and intron 2 of AXIN1 with BMD. Variant_4512 and Variant_4912 encompass the 5' end of their respective genes, thereby implicating the promoter sequence and regulatory elements, which in turn implicates the control of gene expression. Therefore, despite the lack of statistically significant findings in this study, the ESR1 CNV Variant_4512 and the AXIN1 CNV Variant_4912 both still remain as promising candidates for involvement in BMD and the risk of osteoporosis. Moreover, other CNVs in the same genomic regions may also be relevant for future research.

Further research would benefit from addressing the potential effect of environmental risk factors on CNV. It is possible that the ESR1 CNV Variant_4512 may be modified in an environment-specific manner, which influences its effect on BMD, as indicated by the almost statistically significant association between ESR1 CNV and BMD observed in this study when controlled for covariates at the femoral neck ($p = 0.052$). Moreover, previous studies highlight that the majority of known genes subject to CNV are not even located within the identified region of genomic variation, and also that osteoporosis may be more susceptible to genetic variation affecting the CN of non-coding regions. Therefore, further research should also focus on gene expression studies to determine whether the ESR1 CNV Variant_4512 exerts position effects on the transcriptional control of another gene, which may in turn be the primary gene associated with osteoporosis.