



<http://researchspace.auckland.ac.nz>

ResearchSpace@Auckland

Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognise the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

To request permissions please use the Feedback form on our webpage.

<http://researchspace.auckland.ac.nz/feedback>

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the [Library Thesis Consent Form](#) and [Deposit Licence](#).

Note : Masters Theses

The digital copy of a masters thesis is as submitted for examination and contains no corrections. The print copy, usually available in the University Library, may contain corrections made by hand, which have been requested by the supervisor.

***SPRASA*, a novel protein:
An investigation into the
expression profile,
evolutionary conservation
and the association with
infertility.**

Deborah Prendergast

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy
in Obstetrics and Gynaecology
The University of Auckland
2009



Abstract

SPRASA (SPACA3, SLLP1) is a newly identified protein that belongs to the lysozyme super family. Lysozymes are a family of enzymes that are universally expressed in animals, insects and plants and are ubiquitously expressed in a number of tissues and secretions in these organisms. In contrast, *SPRASA* expression appears to be restricted to the equatorial region and the inner acrosomal membrane of sperm. Antibodies reactive with *SPRASA* have been identified in some infertile men and an antiserum reactive with recombinant *SPRASA* prevented human sperm binding to hamster oocytes *in vitro*, indicating an important role in sperm oocyte recognition.

The aim of this study was to investigate the role of *SPRASA* in human infertility. Quantitative real time RT-PCR analysis confirmed the expression of *SPRASA* in the testes, however, in contrast to previous reports, *SPRASA* expression was also identified in the ovaries and heart of both males and females. *In silico* analysis identified two putative promoter regions within the *SPRASA* gene. As expected, the first promoter is located 5' to exon one, while the second promoter is located 5' to exon two. Further investigations by luciferase promoter constructs identified that both promoters were capable of supporting transcriptional activity, however promoter 2 was the more effective promoter. Mutation screening of 102 infertile and 104 fertile couples identified three variants in *SPRASA*. Orthologue sequence analysis indicates that *SPRASA* is a *mammalian-only* gene. This is similar to another member of the lysozyme family, alpha-lactalbumin. However, unlike alpha-lactalbumin, which expresses one variant in all mammals, the *SPRASA* gene's organisation in humans and simians is different to that found in prosimians and non-primates. This suggests that *SPRASA* has a unique function in simians compared to other mammals. Amino acid sequence alignment of the *SPRASA* orthologues revealed a unique motif that is not seen in other lysozyme family members. Computer modelling showed that this *SPRASA* motif is located in a region analogous to the substrate binding region of c-type lysozyme/alpha-lactalbumin. In conclusion, *SPRASA* appears to have a role in both male and female fertility and may be important in species-specific differences in mammalian fertilisation.



Acknowledgements

This work would not have been possible without the advice, assistance, encouragement, guidance and help from my supervisors Andrew Shelling and Larry Chamley, colleagues and friends. My profound thanks to Andrew and Larry for providing much needed guidance, patience and support, without discouraging my efforts. I thank all past and present members of the Shelling and Chamley laboratories for their encouragement, valuable criticism and laughter over the years.

A special mention to my friends Annemarie Newth, Mariann Wichman and Sharon Wichman for all they have done to support me through my years at the University of Auckland. Finally and by no means least, I want to thank Frank, Natasha and Nicholas for their unfailing support and assistance in helping me reach my goals. We made it to the end.

“It is never too late to be what you might have been”

-George Eliot

I would also like to thank the following organisation for funding this research.

The Marsden Fund
Marianne Wichman from ATTACHÉ
NZSAS Trust



Table of Contents

Abstract	ii
Acknowledgements	iii
List of figures	vii
List of tables	ix
Abbreviations	x
Chapter 1: Introduction	2
1 Foreword	2
1.1 Female reproductive system	2
1.2 Male reproductive system.....	6
1.3 Mammalian fertilisation	10
1.3.1 Epididymal maturation	10
1.3.2 Capacitation	11
1.3.3 Vaginal environment	11
1.3.4 Cervical mucus	12
1.3.5 Uterine environment.....	13
1.3.6 The oviductal environment.....	14
1.3.7 Sperm penetration of the cumulus cell layer	14
1.3.8 Sperm-oocyte binding	15
1.3.9 Acrosomal reaction	18
1.3.10 Sperm-oocyte adhesion and fusion	20
1.3.11 Sperm-oocyte adhesion proteins	20
1.3.12 Sperm-oocyte fusion proteins	22
1.4 Infertility	25
1.4.1 Female causes of infertility	26
1.4.2 Ovulation infertility.....	26
1.4.3 Blocked tubes in the female reproductive system	27
1.4.4 Uterine infertility	28
1.4.5 Cervical infertility.....	28
1.4.6 Male causes of infertility	29
1.4.7 Sperm infertility	29
1.4.8 Blocked tubes in the male reproductive system	32
1.4.9 Immunological infertility	32
1.4.10 Genetic abnormalities	33
1.4.11 Combined infertility	35
1.4.12 Unexplained infertility.....	35
1.5 Lysozyme family	35
1.5.1 c-type Lysozyme	35
1.5.2 Alpha-lactalbumin	39
1.5.3 <i>SPRASA</i>	41
1.6 Conclusions from the literature.....	44
1.7 Thesis hypotheses.....	45
Chapter 2: Materials and Methods	47
2 Introduction	47
2.1 Materials	47
2.1.1 Chemicals	47
2.1.2 Enzymes	47
2.1.3 Kits	47
2.1.4 Cell Culture Reagents.....	48
2.1.5 Plasmids	48
2.1.6 Primers.....	48
2.2 Methods	50
2.2.1 Ethics Statement and Patient Recruitment.....	50
2.2.2 DNA Extraction	50
2.2.3 Polymerase Chain Reaction (PCR)	54
2.2.4 Sequencing Reactions	60

2.2.5 Plasmids: Maintaining, Propagation, Isolating and Purification.....	61
2.2.6 Cell culture.....	61
2.2.7 Luciferase Assays.....	62
2.2.8 Total RNA.....	63
2.2.9 Quantitative RT-PCR (qRT-PCT).....	70
2.2.10 Reverse Transcription.....	70
2.2.11 Relative Quantification.....	76
2.2.12 Absolute quantification RT-PCR.....	77
2.2.13 Rapid amplification of cDNA Ends.....	78
2.2.14 <i>In silico</i> analyses.....	79
2.2.15 3D model determination.....	80
2.2.16 Phylogenetic analysis.....	80
2.2.17 Statistical analysis.....	80
Chapter 3: SPRASA expression.....	82
3 Introduction.....	82
3.1 Results.....	82
3.1.1 <i>In silico</i> analysis of <i>SPRASA</i> expression profile.....	84
3.1.2 Quantification real-time PCR.....	88
3.1.3 Spatial and temporal expression analysis of <i>SPRASA</i>	88
3.1.4 Expression analysis of <i>SPRASA</i> in human tissues and cell lines.....	92
3.1.5 Absolute qRT-PCR expression analysis of human <i>SPRASA</i>	94
3.1.6 <i>In silico</i> analysis of <i>SPRASA</i> promoter regions.....	96
3.1.7 Transcription factor binding site analysis.....	98
3.1.8 Functional analysis of <i>SPRASA</i> promoters.....	99
3.2 Discussion.....	101
3.2.1 <i>In silico</i> analysis of the <i>SPRASA</i> protein.....	102
3.2.2 Spatial and temporal expression of <i>SPRASA</i>	104
3.2.3 Expression pattern of <i>SPRASA</i> mRNA in the male mouse reproductive system.....	105
3.2.4 Human <i>SPRASA</i> expression.....	106
3.2.5 Absolute expression of <i>SPRASA</i> in human samples.....	107
3.2.6 <i>SPRASA</i> promoter analysis.....	108
3.2.7 Analysis of the promoter regions of <i>SPRASA</i>	109
3.3 Summary.....	112
Chapter 4: Mutational analysis of the SPRASA gene.....	114
4 Introduction.....	114
4.1 Results.....	116
4.1.1 Optimisation of PCR amplification.....	116
4.1.2 Calculation of allele, homozygote and heterozygote frequencies.....	117
4.1.3 Nucleotide analysis of exon one.....	118
4.1.4 Functional analysis of the <i>SPRASA</i> promoter variant g.-22TGC(4_5).....	121
4.1.5 Nucleotide analysis of exon two.....	122
4.1.6 <i>In silico</i> analysis of c.314G>A (p.C80Y) variant.....	127
4.1.7 Nucleotide analysis of exon three and exon four.....	130
4.1.8 Nucleotide analysis of exon five.....	130
4.1.9 Variant analysis within couples.....	132
4.2 Discussion.....	135
4.2.1 Tri-nucleotide insertion at the quadruple tri-nucleotide repeat region.....	135
4.2.2 Luciferase assay characterisation of the <i>SPRASA</i> g.-22TGC(4_5) variant.....	138
4.2.3 The <i>SPRASA</i> c.314G>A possible association with infertility.....	139
4.2.4 Genetic analyses of exon three and exon four of the <i>SPRASA</i> gene.....	143
4.2.5 Novel nucleotide change in the 3' UTR of the <i>SPRASA</i> gene.....	144
4.2.6 Variation analysis within infertile and fertile couples.....	144
4.3 Summary.....	146
Chapter 5: Evolution of the SPRASA gene.....	148
5 Introduction.....	148
5.1 Results.....	149
5.1.1 <i>SPRASA</i> orthologues.....	149
5.1.2 Gene organisation and structural homology.....	161



5.1.3 <i>SPRASA</i> motif.....	169
5.1.4 Phylogenetic analysis of amino acid sequence data.....	171
5.1.5 Tertiary structure modelling.....	174
5.2 Discussion.....	175
5.2.1 <i>SPRASA</i> orthologue analysis.....	175
5.2.2 Analysis of the orang-utan <i>SPRASA</i> gene.....	176
5.2.3 <i>SPRASA</i> gene organisation and structural homology.....	180
5.2.4 Catalytic residues.....	181
5.2.5 Analysis of <i>SPRASA</i> invariant residues.....	182
5.2.6 Analysis of the <i>SPRASA</i> motif.....	183
5.2.7 Phylogenetic analysis.....	184
5.2.8 Analysis of the tertiary structure modelling.....	186
5.3 Summary.....	188
Chapter 6: Concluding remarks.....	190
6 Discussion.....	190
6.1 The determination of the spatial and temporal expression of <i>SPRASA</i>	190
6.2 To determine whether genetic variants in the <i>SPRASA</i> gene sequence are associated with infertility in humans.....	193
6.3 To examine the evolutionary conservation of <i>SPRASA</i>	196
6.4 Future directions.....	198
6.5 Summary.....	200
Appendix I.....	202
7 <i>SPRASA</i> promoter and gene sequence.....	202
Appendix II.....	206
8 Buffer and Solutions.....	206
8.1 Stock solutions for DNA isolation.....	206
8.2 Buffers and solutions for gel electrophoresis.....	209
8.3 RNA visualisation on agarose gel.....	212
8.4 Cell transformation.....	212
8.5 General solutions.....	213
8.6 Sequencing cleanup.....	214
8.7 Cell Culture.....	215
8.8 Competent Cells Solutions.....	216
Appendix III.....	218
References.....	235



List of figures

Chapter 1

Figure 1. 1: Female organs of reproduction.	4
Figure 1. 2: The human ovary.	5
Figure 1. 3: Diagrammatic representation of an ovulated oocyte.....	6
Figure 1. 4: Male testes and associated structures.....	7
Figure 1. 5: Diagrammatic representation of a human sperm.....	8
Figure 1. 6: Diagrammatic representation of mammalian sperm acrosome and head morphology.....	9
Figure 1. 7: Biological basis for human capacitation.....	12
Figure 1. 8: Mammalian fertilisation.....	15
Figure 1.9: The acrosomal reaction.....	19
Figure 1. 10: The location of <i>SPRASA</i> on human sperm.....	22
Figure 1. 11: Various abnormal sperm morphology.....	31
Figure 1. 12: Structure of the human c-type lysozyme gene.....	36
Figure 1. 13: Structure of the human c-type lysozyme protein.....	37
Figure 1. 14: Three dimensional image of the mature human c-type lysozyme.....	39
Figure 1. 15: Structure of the human <i>SPRASA</i> gene.....	42
Figure 1. 16: Structure of the human <i>SPRASA</i> protein.....	42

Chapter 2

Figure 2. 1: Oragene•DNA/saliva Self Collection Kit.....	52
Figure 2. 2: Human <i>SPRASA</i> Genomic Sequence.....	54
Figure 2. 3: Primers for promoter variant amplification and sequencing.....	55
Figure 2. 4: Diagrammatic representation of the section taken for quantitative real-time RT-PCR of an adult human testis showing general structure of the tissue.....	64
Figure 2. 5: Human real time amplicon design.....	72
Figure 2. 6: Mouse real time amplicon design.....	72

Chapter 3

Figure 3. 1: Similarity between the mouse, rat and human <i>SPRASA</i> amino acid sequences and the human c-type lysozyme.....	84
Figure 3. 2: Graphical outputs of SignalP and Phobius prediction.....	87
Figure 3. 3: Normalised spatial and temporal expression of <i>SPRASA</i> in female and male C57 mice at the 20 days postnatal time point.....	89
Figure 3. 4: Alignment and orientation of the mouse quantitative RT-PCR primers and probe.....	90
Figure 3. 5: Expression analysis of <i>SPRASA</i> mRNA in the CD1 mice.....	91
Figure 3. 6: Expression analysis of <i>SPRASA</i> mRNA in fertile and vasectomised 1 year old CD1 male mice.....	91
Figure 3. 7: Expression analysis of <i>SPRASA</i> isoforms in human tissues and cell lines.....	93
Figure 3. 8: Absolute expression of the short (<i>SPRASA</i> 4/5) isoform of <i>SPRASA</i> in human tissues and cell lines expressing <i>SPRASA</i>	95
Figure 3. 9: Cister promoter prediction of the <i>SPRASA</i> gene.....	97
Figure 3. 10: Genetic organisation of the human <i>SPRASA</i> promoter 2.....	98
Figure 3. 11: Functional analysis of the two human putative promoter regions by luciferase reporter assay.....	100

Chapter 4

Figure 4. 1: Agarose gel images of PCR amplicons and promoter amplicons.....	117
Figure 4. 2: Variant expressing exon amplicons.....	118
Figure 4. 3: Tri-nucleotide repeat polymorphism found in exon 1 of the <i>SPRASA</i> gene.....	119
Figure 4. 4: Functional analysis of the two human promoter 1 variants by luciferase reporter assay.....	122
Figure 4. 5: Non-synonymous amino acid change in the transmembrane region of exon 2.....	122
Figure 4. 6: The graphical outputs of TMHMM2.0 and SignalP.....	128
Figure 4. 7: Cartoon representation of the 2D image of the mature <i>SPRASA</i> protein.....	130
Figure 4. 8: G to C transversion in the 3' UTR of the <i>SPRASA</i> gene.....	132
Figure 4. 9: Non-B DNA conformation.....	137



Figure 4. 10: Signal peptide structure and signal anchor structure..... 141

Chapter 5

Figure 5. 1: *SPRASA* multiple sequence alignment..... 149
Figure 5. 2: Alignment of the amino acid sequence of *SPRASA* orthologues. 154
Figure 5. 3: Non-synonymous amino acid change in exon two of the orang-utan *SPRASA* gene..... 157
Figure 5. 4: Comparison of c-type lysozyme and *SPRASA* exon organisation..... 162
Figure 5. 5: Cartoon representation of the 2D image of the long isoform of the *SPRASA* protein..... 171
Figure 5. 6: Phylogenetic tree of *SPRASA* orthologues..... 172
Figure 5. 7: 3D comparison view of the mature c-type lysozyme and *SPRASA* proteins. 174
Figure 5. 8: Comparison of the Kozak sequence observed surrounding the two putative start sites. 178
Figure 5. 9: Catalytic residues of the c-type lysozyme family. 182
Figure 5. 10: Proposed scheme for the evolution of *SPRASA*..... 186

Chapter 6

Figure 6. 1: 2D structure similarity of c-type lysozyme, alpha-lactalbumin and *SPRASA* family. 197

Appendix III

Figure 1: Real time RT-PCR amplification plote of human *SPRASA* expression in the testis and ovary..... 218
Figure 2: Function analysis of *SPRASA* promoters..... 221
Figure 3: Function analysis of *SPRASA* promoter variants..... 222
Figure 4: Venn diagrams of infertile individuals / couples. 223
Figure 5: Venn diagrams of fertile individuals / couples..... 224
Figure 6: Dog *SPRASA* mRNA sequence 225
Figure 7: Chimpanzee *SPRASA* DNA sequence 226
Figure 8: Orang-utan *SPRASA* DNA sequence 228
Figure 9: Baboon *SPRASA* DNA sequence 230
Figure 10: Capuchin monkey *SPRASA* DNA sequence 232



List of tables

Chapter 1

Table 1. 1: Phenotype of oocyte and sperm from knockout mice in sperm-oocyte interaction.....	21
---	----

Chapter 2

Table 2. 1: Orthologue DNA extraction protocols.....	52
Table 2. 2: Summary of primers used for mutation screening and promoter amplification.....	56
Table 2. 3: Summary of primers used for <i>SPRASA</i> exon amplification of orthologues.	59
Table 2. 4: Separation of DNA fragments on agarose gels.....	59
Table 2. 5: General information on total human and mouse RNA and cDNA from human cell lines.	69
Table 2. 6: Summary of primer pairs and probes used for human qRT-PCR analyses.....	74
Table 2. 7: Summary of primers and probe used for mouse qRT-PCR analyses.....	75

Chapter 3

Table 3. 1: Calibrated <i>SPRASA</i> expression of the short (<i>SPRASA</i> 4/5) and long (<i>SPRASA</i> 1/2) isoforms of <i>SPRASA</i> in the human tissues and cell lines expressing <i>SPRASA</i>	94
Table 3. 2: Calibrated <i>SPRASA</i> expression.	96

Chapter 4

Table 4. 1: Summary of aetiologies of infertility.....	115
Table 4. 2: Population frequencies of the tri-nucleotide repeat (g.-22TGC(4_5)).	120
Table 4. 3: Population frequencies of the G>A transition (c.314G>A).	126
Table 4. 4: In silico prediction of the c.314G>A (p.C80Y) transition in exon 2.....	127
Table 4. 5: Population frequencies of the G>C transversion (c.766G>C).....	131
Table 4. 6: Variant frequencies of g.-22TGC(4_5) and c.314G>A in fertile and infertile couples.	134

Chapter 5

Table 5. 1: Information on <i>SPRASA</i> orthologues.	160
Table 5. 2: Evolutionary conservation of exon-intron junctions in the <i>SPRASA</i> orthologues.	163
Table 5. 3: Homology of <i>SPRASA</i> orthologues.....	165
Table 5. 4: Exon conservation to the human <i>SPRASA</i> protein.....	166
Table 5. 5: Amino acid residues observed at each position among <i>SPRASA</i> orthologues..	168
Table 5. 6: Comparison of the <i>SPRASA</i> motif in 33 orthologues.....	169

Appendix III

Table 1: Calculation for the ovarian transcript copy number per ng.....	219
Table 2: Calibrated <i>SPRASA</i> expression. Datum is presented as absolute copy number..	220



Abbreviations

α	Alpha
ACTB	Actin, beta
AR	Androgen receptor
ATCC	American Tissue Culture Collection
B2M	Beta-2 microglobulin
bp	Base pair
cDNA	Complementary DNA
C.O.D	Cause of death
C _T	Threshold cycle
Cister	<u>Cis-elements clusters</u>
CYC1	Cytochrome C-1
dNTPs	Deoxyribonucleic acids
ΔR_n	Delta reaction ($\Delta R_n = R_{n+} - R_{n-}$)
EDTA	Ethylenediaminetetraacetic acid
FCS	Fetal calf serum
GFP	Green fluorescent protein
GMS	Grantham matrix score
GR	Glucocorticoid receptor
HEK-293T	Human embryonic kidney 293T
HPRT1	Hypoxanthine guanine phosphoribosyltransferase 1
Hrs	Hours
Kb	Kilobase
KDa	Kilodalton
LB	Luria bertani
Luc/lux	Luciferase
mRNA	messenger RNA
myr	million years
NCBI	National Centre for Biotechnology Information
NF	Normalisation factor
P20	Perinatal, 20 days old
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PK	Proteinase K
PolyPhen	Polymorphism phenotyping
PPIA	Peptidyl-proly isomerise A
PR	Progesterone receptor
qRT-PCR	Quantitative reverse transcription PCR
RACE	Rapid amplification of cDNA ends
Rn+	Fluorescence emission of the product at each time point
Rn-	Fluorescence emission of the baseline
RNA	Ribonucleic acid
RPL13A	Ribosomal protein L13a



RPLP0	Ribosomal phosphoprotein large P0
Rpm	revolutions per minute
RPMI 1640	Roswell Park Memorial Institute 1640
RT	Reverse transcriptase
SDHA	Succinate dehydrogenase complex, subunit A
SDS	Sodium dodecyl sulphate
SIFT	Sorts intolerant from tolerant
SNP	Single nucleotide polymorphism
SPRASA	<u>S</u> perm <u>P</u> rotein <u>R</u> eactive with <u>A</u> nti <u>S</u> perm <u>A</u> ntibodies
TAE	Tris acetic acid EDTA
TBP	TATA box binding protein
TESS	Transcription element search software
TE	Tris-EDTA
TF	Transcription factor
TMHMM2.0	<u>T</u> rans <u>m</u> embrane <u>h</u> idden <u>M</u> arkov <u>m</u> odels version 2.0
UBC	Ubiquitin C
UTR	Untranslated region
UV	Ultraviolet
v/v	volume per volume
w/v	Weight per volume
YWHAZ	Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide