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](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
Chapter 3: SPRASA expression ........................................... 82

3 Introduction........................................................................ 82

3.1 Results ........................................................................... 82

3.1.1 In silico analysis of SPRASA expression profile .......... 84
3.1.2 Quantification real-time PCR ....................................... 88
3.1.3 Spatial and temporal expression analysis of SPRASA .... 88
3.1.4 Expression analysis of SPRASA in human tissues and cell lines ................................................................. 92
3.1.5 Absolute qRT-PCR expression analysis of human SPRASA ............................................................... 94
3.1.6 In silico analysis of SPRASA promoter regions .......... 96
3.1.7 Transcription factor binding site analysis .................... 98
3.1.8 Functional analysis of SPRASA promoters .................. 99

3.2 Discussion ..................................................................... 101

3.2.1 In silico analysis of the SPRASA protein ....................... 102
3.2.2 Spatial and temporal expression of SPRASA ................. 104
3.2.3 Expression pattern of SPRASA mRNA in the male mouse reproductive system 105
3.2.4 Human SPRASA expression ...................................... 106
3.2.5 Absolute expression of SPRASA in human samples ....... 107
3.2.6 SPRASA promoter analysis ....................................... 108
3.2.7 Analysis of the promoter regions of SPRASA .......... 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 SPRASA promoter variant g-22TGC(4_5) ..................................................... 121
4.1.5 Nucleotide analysis of exon two .................................. 122
4.1.6 In silico 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 SPRASA g-22TGC(4_5) variant ................. 138
4.2.3 The SPRASA c.314G>A possible association with infertility ................................................................. 139
4.2.4 Genetic analyses of exon three and exon four of the SPRASA gene ................................. 143
4.2.5 Novel nucleotide change in the 3’ UTR of the SPRASA 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 SPRASA orthologues .................................................. 149
5.1.2 Gene organisation and structural homology ................. 161
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 SPRASA 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 SPRASA gene. ................................................ 42
Figure 1. 16: Structure of the human SPRASA protein. ............................................. 42

Chapter 2
Figure 2. 1: Oragene•DNA/saliva Self Collection Kit. ................................................ 52
Figure 2. 2: Human SPRASA 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 SPRASA 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 SPRASA 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 SPRASA mRNA in the CD1 mice. ....................... 91
Figure 3. 6: Expression analysis of SPRASA mRNA in fertile and vasectomised 1 year old CD1 male mice. ......................................................................................... 91
Figure 3. 7: Expression analysis of SPRASA isoforms in human tissues and cell lines. .... 93
Figure 3. 8: Absolute expression of the short (SPRASA 4/5) isoform of SPRASA in human tissues and cell lines expressing SPRASA. ............................................. 95
Figure 3. 9: Cister promoter prediction of the SPRASA gene. .................................... 97
Figure 3. 10: Genetic organisation of the human SPRASA 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 SPRASA 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 SPRASA protein... 130
Figure 4. 8: G to C transversion in the 3’ UTR of the SPRASA gene. ......................... 132
Figure 4. 9: Non-B DNA conformation...................................................................... 137
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 plot 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 SPRASA 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 SPRASA expression of the short (SPRASA 4/5) and long (SPRASA 1/2) isoforms of SPRASA in the human tissues and cell lines expressing SPRASA........... 94
Table 3. 2: Calibrated SPRASA 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 SPRASA orthologues. .............................................................. 160
Table 5. 2: Evolutionary conservation of exon-intron junctions in the SPRASA orthologues. ............................................................................................................................... 163
Table 5. 3: Homology of SPRASA orthologues. ............................................................... 165
Table 5. 4: Exon conservation to the human SPRASA protein. ........................................ 166
Table 5. 5: Amino acid residues observed at each position among SPRASA orthologues. 168
Table 5. 6: Comparison of the SPRASA motif in 33 orthologues. ................................. 169

Appendix III
Table 1: Calculation for the ovarian transcript copy number per ng.............................. 219
Table 2: Calibrated SPRASA expression. Datum is presented as absolute copy number. 220
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>Alpha</td>
</tr>
<tr>
<td>ACTB</td>
<td>Actin, beta</td>
</tr>
<tr>
<td>AR</td>
<td>Androgen receptor</td>
</tr>
<tr>
<td>ATCC</td>
<td>American Tissue Culture Collection</td>
</tr>
<tr>
<td>B2M</td>
<td>Beta-2 microglobulin</td>
</tr>
<tr>
<td>bp</td>
<td>Base pair</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary DNA</td>
</tr>
<tr>
<td>C.O.D</td>
<td>Cause of death</td>
</tr>
<tr>
<td>Ct</td>
<td>Threshold cycle</td>
</tr>
<tr>
<td>Cister</td>
<td>Cis-elements clusters</td>
</tr>
<tr>
<td>CYC1</td>
<td>Cytochrome C-1</td>
</tr>
<tr>
<td>dNTPs</td>
<td>Deoxyribonucleic acids</td>
</tr>
<tr>
<td>ΔRn</td>
<td>Delta reaction (ΔRn=Rn+ - Rn-)</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>FCS</td>
<td>Fetal calf serum</td>
</tr>
<tr>
<td>GFP</td>
<td>Green fluorescent protein</td>
</tr>
<tr>
<td>GMS</td>
<td>Grantham matrix score</td>
</tr>
<tr>
<td>GR</td>
<td>Glucocorticoid receptor</td>
</tr>
<tr>
<td>HEK-293T</td>
<td>Human embryonic kidney 293T</td>
</tr>
<tr>
<td>HPRT1</td>
<td>Hypoxanthine guanine phosphoribosyltransferase 1</td>
</tr>
<tr>
<td>Hrs</td>
<td>Hours</td>
</tr>
<tr>
<td>Kb</td>
<td>Kilobase</td>
</tr>
<tr>
<td>KDa</td>
<td>Kilodalton</td>
</tr>
<tr>
<td>LB</td>
<td>Luria bertani</td>
</tr>
<tr>
<td>Luc/lux</td>
<td>Luciferase</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>myr</td>
<td>million years</td>
</tr>
<tr>
<td>NCBI</td>
<td>National Centre for Biotechnology Information</td>
</tr>
<tr>
<td>NF</td>
<td>Normalisation factor</td>
</tr>
<tr>
<td>P20</td>
<td>Perinatal, 20 days old</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PK</td>
<td>Proteinase K</td>
</tr>
<tr>
<td>PolyPhen</td>
<td>Polymorphism phenotyping</td>
</tr>
<tr>
<td>PPIA</td>
<td>Peptidyl-proly isomerise A</td>
</tr>
<tr>
<td>PR</td>
<td>Progesterone receptor</td>
</tr>
<tr>
<td>qRT-PCR</td>
<td>Quantitative reverse transcription PCR</td>
</tr>
<tr>
<td>RACE</td>
<td>Rapid amplification of cDNA ends</td>
</tr>
<tr>
<td>Rn+</td>
<td>Fluorescence emission of the product at each time point</td>
</tr>
<tr>
<td>Rn-</td>
<td>Fluorescence emission of the baseline</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RPL13A</td>
<td>Ribosomal protein L13a</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>RPLP0</td>
<td>Ribosomal phosphoprotein large P0</td>
</tr>
<tr>
<td>Rpm</td>
<td>revolutions per minute</td>
</tr>
<tr>
<td>RPMI 1640</td>
<td>Roswell Park Memorial Institute 1640</td>
</tr>
<tr>
<td>RT</td>
<td>Reverse transcriptase</td>
</tr>
<tr>
<td>SDHA</td>
<td>Succinate dehydrogenase complex, subunit A</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulphate</td>
</tr>
<tr>
<td>SIFT</td>
<td>Sorts intolerant from tolerant</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
<tr>
<td>SIFT</td>
<td>Sperm Protein Reactive with AntiSperm Antibodies</td>
</tr>
<tr>
<td>TAE</td>
<td>Tris acetic acid EDTA</td>
</tr>
<tr>
<td>TBP</td>
<td>TATA box binding protein</td>
</tr>
<tr>
<td>TESS</td>
<td>Transcription element search software</td>
</tr>
<tr>
<td>TE</td>
<td>Tris-EDTA</td>
</tr>
<tr>
<td>TF</td>
<td>Transcription factor</td>
</tr>
<tr>
<td>TMHMM2.0</td>
<td>Transmembrane hidden Markov models version 2.0</td>
</tr>
<tr>
<td>UBC</td>
<td>Ubiquitin C</td>
</tr>
<tr>
<td>UTR</td>
<td>Untranslated region</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>v/v</td>
<td>volume per volume</td>
</tr>
<tr>
<td>w/v</td>
<td>Weight per volume</td>
</tr>
<tr>
<td>YWHAZ</td>
<td>Tyrosine 3-monoxygenase/tryptophan 5-monoxygenase activation protein, zeta polypeptide</td>
</tr>
</tbody>
</table>