# TABLE OF CONTENTS

Abbreviations ....................................................................................... i-iii

1. INTRODUCTION AND OBJECTIVES .............................................. 1-4

2. REVIEW OF LITERATURE ................................................................. 5-23

2.1 *Bubalus bubalis*: An important livestock species ....................... 5-6
2.2 Repetitive Sequences........................................................................ 6
  2.2.1 Satellites ................................................................................. 7-8
    2.2.1.1 Satellites in the non-coding genomes ................................. 8
    2.2.1.2 Satellite DNA transcripts and their implications .......... 8-9
  2.2.2 Simple Sequence Repeats (SSRs) ....................... 9-10
    2.2.2.1 Evolution of SSRs ......................................................... 10-11
    2.2.2.2 Characteristics of SSRs ............................................. 11-12
    2.2.2.3 Distribution of SSRs ................................................ 12-14
    2.2.2.4 Putative functions and effects of SSRs in genome ........ 14-20
      2.2.2.4.1 In the Non-coding Genome ................................. 14-18
      2.2.2.4.2 In the Coding Genome .................................. 18-20
    2.2.2.5 SSRs and diseases linkage ..................................... 20-21
  2.3 Minisatellite/Microsatellite Associated Sequence Amplification .... 21-23

3. MATERIALS AND METHODS ............................................................. 24-33

3.1 Sample collection and genomic DNA isolation .............................. 24
3.2 Isolation of total RNA and cDNA synthesis ............................... 25
3.3 Minisatellite/Microsatellite Associated Sequence Amplification ...... 26
3.4 Cloning, Sequencing and Characterization of the MASA uncovered Amplicons ................................................ 26-27
3.5 Cloning and characterization of *c-kit* and *Smoc-1* genes ........... 27-28
3.6 Homology status, phylogenetic delineation, domain organization, secondary structure prediction and other in-silico analyses .... 28
3.7 RNA slot blot analysis and Northern blotting .............................. 28-29
3.8 RT-PCR and Southern Blotting .................................................. 29
3.9 Evolutionary studies of the uncovered genes/gene fragments ........ 30
3.10 Copy number calculation and Relative expressional studies .......... 31
3.11 Metaphase chromosome preparation and Fluorescent in situ hybridization .......................... 31-32

3.12 Protein expression and production of anti-Smoc-1 antiserum ........... 32-33

3.13 Isolation of total protein from tissues and Western Blotting .......... 33

3.14 Immunohistochemical analysis of Smoc-1 on Tissue Sections ......33

4. RESULTS ........................................................................................................................................ 34-58

4.1 MASA mediated mining of the mRNA transcripts ....................... 34-46

4.1.1 Distribution of GACA/GATA/33.15 repeats .......................... 34-35

4.1.1.1 Repeat motifs within the non-coding genomes .................. 34-35

4.1.1.2 Repeat motifs within the coding genomes ...................... 35

4.1.2 Identification of the mRNA transcripts by MASA, followed by detailed Characterization .......................................................................................................................... 35-40

4.1.2.1 MASA with consensus sequence of 33.15 repeat ............ 36-37

4.1.2.2 Simple repeat of GACA uncovered more number of transcripts in spermatozoa ........................................... 37-38

4.1.2.3 Transcripts identified by GATA repeat in different tissues and spermatozoa ......................................................... 38-39

4.1.2.4 Comparative analysis of the transcript diversity ............ 39-40

4.1.3 Sequence polymorphisms in the MASA uncovered transcripts ...40-42

4.1.3.1 Within the 33.15 tagged transcripts .............................. 40-41

4.1.3.2 In the GACA tagged transcripts .................................. 41

4.1.3.3 Within the GATA tagged transcripts ............................. 42

4.1.4 Conservation of the entrapped genes across the species ...... 42

4.1.5 Copy number status of the uncovered genes ...................... 43

4.1.6 Differential expression of the repeat tagged genes .................. 43-46

4.1.6.1 Of the 33.15 repeat tagged genes ................................ 44-45

4.1.6.2 Of the GACA tagged transcripts ................................. 45

4.1.6.3 Of the GATA tagged transcripts .................................. 46

4.1.7 Chromosomal mapping ................................................. 46

4.2 Isolation and detailed characterization of candidate genes ...... 47-58

4.2.1 Proto-oncogene C-kit receptor ........................................ 47-51

4.2.1.1 Isolation of full length CDS of buffalo c-kit receptor ........ 47-48

4.2.1.2 Domain organization of buffalo c-kit receptor ................ 48
4.2.1.3 Tissue specific sequence variations ........................................... 48-49
4.2.1.4 Alternate splicing of buffalo c-kit in different tissues .......... 49
4.2.1.5 Uniqueness in tyrosine kinase domain of buffalo c-kit ....... 49-50
4.2.1.6 Evolutionary relationship of buffalo c-kit gene ...................... 50
4.2.1.7 Buffalo genome has single copy of c-kit gene ....................... 51
4.2.1.8 Relative expression of c-kit receptor gene .......................... 51
4.2.2 Secreted Modular Calcium Binding Protein-1 ......................... 52-58
4.2.2.1 Isolation and cloning of buffalo Smoc-1 gene..................... 52
4.2.2.2 Buffalo Smoc-1 shows two transcript variants ................... 53
4.2.2.3 Structure of Smoc-1 and phylogenetic delineation ............ 53-54
4.2.2.4 Domain Organization of Smoc-1 ........................................... 54-55
4.2.2.5 Single copy of the Smoc-1 gene located on chromosome 11 in
the Buffalo .................................................................................. 55
4.2.2.6 Recombinant expression of Smoc-1 .................................. 55-56
4.2.2.7 Highest expression of Smoc-1 transcript variants in liver.... 56-57
4.2.2.8 Age specific expression profile of the Smoc-1 ............... 57
4.2.2.9 Association of Smoc-1 with the basement membrane ..... 57-58

5. DISCUSSION .................................................................................. 59-72

5.1 Repeat tagged transcript diversity ......................................... 59
5.1.1 Global distribution of the 33.15, GACA and GATA repeats ..... 59-60
5.1.2 Biological significance of the repeats within the somatic and
spermatozoal transcriptomes ....................................................... 61
5.1.3 Implications of organizational variations uncovered .......... 62-63
5.1.4 Prospects of repeat tagged transcripts carrying highest expression in
the testis and spermatozoa ......................................................... 64-65
5.1.5 MASA and comparative genomics ....................................... 65-66

5.2 Differential organization and expression of Smoc-1 and C-kit in
Water buffalo ............................................................................... 66-72
5.2.1 C-kit and its tissue specific nature ...................................... 66-68
5.2.1.1 C-kit: structure and domain organization ...................... 66-67
5.2.1.2 Tissue specific alternate splicing and c-kit .................... 67
5.2.1.3 C-kit with testis specific expression .............................. 67-68
5.2.2 Smoc-1 and its transcript variants ...................................... 68-72
5.2.2.1 Smoc-1 and SPARC family: Comparative organization ..... 68-69
5.2.2.2 Potential implications of multi-domain proteins .................. 69
5.2.2.3 Smoc-1 and its transcript variants in different species ........... 69-70
5.2.2.4 Single copy Smoc-1 showed highest expression in liver .... 70-71
5.2.2.5 Smoc-1 and age specific expression ............................... 71
5.2.2.6 Smoc-1: tissue localization and future aspects .................. 71-72

6. CONCLUSIONS ............................................................................. 73-76

7. REFERENCES .............................................................................. 77-97

8. LIST OF PUBLICATIONS FROM THIS STUDY .............................. 98