From: the department of molecular medicine and surgery
Karolinska Institutet, Stockholm, Sweden

THE TESTATIN/CRES SUBGROUP OF FAMILY 2
CYSTATINS IN SEXUAL DEVELOPMENT

Jessica Frygelius

Stockholm 2009
Cover illustration: Olle Bærtling, Iru, 1958, olja på duk. Moderna Museet

All previously published papers were reproduced with permission from the publisher.

Published by Karolinska Institutet. Printed by Larseric Digital print AB.
© Jessica Frygelius, 2009
ABSTRACT

The genetic sex in mammals is established at fertilization, where an XX chromosome complement will lead to female development and an XY genotype will result in male development. Thereafter, sex determination is initiated with expression of the Y-linked factor Sry in XY gonads, leading to testis formation, while in the absence of Sry expression an ovary is formed. A number of genes, both transcription factors and signaling molecules such as Sox9, Fgf9, Dhh, Dax1 and Wnt4 have been identified downstream of Sry. Disruption of these genes causes sex reversal and/or impaired gonad development in humans and/or mouse models. However, it is clear that key factors in gonad differentiation remain to be identified. Finding new genes would not only clarify basic mechanisms behind gonad development but could also improve diagnostics in patients with disorders of sex development.

Testatin was previously isolated by our group in a screen searching for novel genes expressed in early mouse sex differentiation. Testatin is specifically up regulated in the developing testis just after expression of Sry. Testatin belongs to the Testatin/Cres subgroup of cystatin family 2 protease inhibitors that show a reproductive tract restricted expression (testis, epididymis, ovary, pituitary) in contrast to the broad expression profiles of classical family 2 cystatins, implying specialized functions in reproduction.

To evaluate the role of Testatin in male sexual development, we generated a Testatin knockout mouse. Detailed phenotyping revealed normal testis development and fertility in male Testatin knockout animals (Paper I). An explanation for the lack of phenotype in the knockout mice could be functional redundancy between the subgroup members. Therefore, we evaluated the expression profiles of the Testatin/Cres subgroup genes in fetal testis using real-time PCR and in situ hybridization. We show that three of the subgroup members, namely Cres, cystatin SC, Cystatin TE-1, are expressed in mouse fetal testis together with Testatin (Paper II).

Ancestors of the Cystatins can be traced back to plants and to understand when and why the Testatin/Cres subgroup genes emerged, we performed an evolutionary study (Paper III). We have localized the evolutionary origin of the Testatin/Cres subgroup genes to the split between Marsupials and placental mammals and a model for the evolution of these genes illustrates that they constitute a dynamic group of genes, which has undergone several gene expansions. Further, indications of a high degree of positive selection, in striking contrast to what is seen for the classical Cystatin C was found. We suggest a new nomenclature for the Testatin/Cres subgroup (TCS) genes, based on their syntenic gene location on mouse chromosome 2 and show with phylogenetic relations that the TCS genes are clustered into three original groups, a testatin (TCS7), a Cres (TCS5) and a CstL1 (TCS1) group.

We also evaluated the expression patterns of all human members of the subfamily (Paper III). Of a total of nine identified human genes, four express putative functional transcripts with a predominant expression in the male reproductive system which is in line with a suggested role in reproduction.

In conclusion, the expression profiles and the evolutionary history of the Testatin/Cres subgroup genes are compatible with a role of these genes in reproduction in placental mammals. Future studies, such as the generation of mice with targeted deletion encompassing additional TCS genes, could finally resolve this issue.
LIST OF PUBLICATIONS


TABLE OF CONTENTS

INTRODUCTION.................................................................................................................. 1
  Evolution of animal species.......................................................................................... 1
  Sexual development....................................................................................................... 2
  Sexual development in mammals................................................................................ 2
  Sex determination in eutherian mammals............................................................... 5
  Mouse testis formation.................................................................................................. 7
  Molecular interactions in mouse testis differentiation............................................. 9
  Disorders of sexual development (DSD) in XY individuals.................................... 11
  Testatin......................................................................................................................... 12
  Cystatin super family.................................................................................................. 12
  Family 2 cystatins........................................................................................................ 12
  Testatin/Cres subgroup of family 2 cystatins........................................................... 13
AIMS.................................................................................................................................. 17

METHODS AND MATERIALS............................................................................................. 18
  Methods in unpublished project A............................................................................... 18
  Methods in unpublished project B............................................................................... 19
  Methods in unpublished project C............................................................................... 20

RESULTS.......................................................................................................................... 22
  Normal sexual development and fertility in testatin knockout mice (Paper I)........... 22
  Body weight in testatin deficient mice (unpublished project A)................................. 22
  The reproductive tissue specific cystatin subgroup of genes: expression during
  gonadal development in wild type and testatin knockout animals (Paper II)............ 24
  Evolution and human tissue expression of the TCS genes, a reproductive tissue
  specific subgroup of the family 2 cystatins (Paper III)............................................... 24
  The Testatin/Cres subgroup genes in disorders of sexual development
  (unpublished project B)............................................................................................... 26
  Production of recombinant proteins (unpublished project C).................................... 26

DISCUSSION....................................................................................................................... 28
  Putative functions of the TCS genes in testis development........................................ 28
  Other putative functions of the TCS genes............................................................... 30
  Spatial localization of the TCS proteins..................................................................... 32

CONCLUSIONS AND FUTURE DIRECTIONS................................................................... 33

ACKNOWLEDGEMENTS................................................................................................... 34

REFERENCES.................................................................................................................... 36
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>aa</td>
<td>amino acids</td>
</tr>
<tr>
<td>AMH</td>
<td>Anti-Müllerian hormone</td>
</tr>
<tr>
<td>Cres</td>
<td>Cystatin-related epididymal spermatogenic protein/gene</td>
</tr>
<tr>
<td>CLM</td>
<td>Cystatin like molecule</td>
</tr>
<tr>
<td>Cst/CST</td>
<td>Cystatin</td>
</tr>
<tr>
<td>CstT</td>
<td>Testis specific cystatin like</td>
</tr>
<tr>
<td>CstL1/9</td>
<td>Cystatin like gene 1/9</td>
</tr>
<tr>
<td>CstTE1</td>
<td>Cystatin related gene highly expressed in testis and epididymis</td>
</tr>
<tr>
<td>CstSC</td>
<td>Cystatin related gene expressed in Sertoli cells</td>
</tr>
<tr>
<td>DAX1</td>
<td>DSS Adrenal Hypoplasia Congenital critical region on chromosome X, gene 1</td>
</tr>
<tr>
<td>DEXA</td>
<td>Dual energy X-ray absorptiometry</td>
</tr>
<tr>
<td>Dhh</td>
<td>Desert hedgehog homologue</td>
</tr>
<tr>
<td>DMRT1</td>
<td>Doublesex and Mab-3 related transcription factor 1</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DSD</td>
<td>Disorders of sexual development</td>
</tr>
<tr>
<td>DSS</td>
<td>Dosage sensitive sex reversal</td>
</tr>
<tr>
<td>E</td>
<td>Embryonic day</td>
</tr>
<tr>
<td>Fgf9</td>
<td>Fibroblast growth factor 9</td>
</tr>
<tr>
<td>Fgfr2</td>
<td>Fibroblast growth factor receptor 2</td>
</tr>
<tr>
<td>FSH</td>
<td>Follicle stimulating hormone</td>
</tr>
<tr>
<td>FOXL2</td>
<td>Forkhead/winged helix transcription factor 2</td>
</tr>
<tr>
<td>Ir</td>
<td>Insulin receptor</td>
</tr>
<tr>
<td>Igf1r</td>
<td>Insulin-like growth factor 1 receptor</td>
</tr>
<tr>
<td>Irr</td>
<td>Insulin-related receptor</td>
</tr>
<tr>
<td>kDa</td>
<td>kiloDalton</td>
</tr>
<tr>
<td>LH</td>
<td>Luteinising hormone</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger ribonucleic acid</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>Pdgfα</td>
<td>Platelet-derived growth factor α</td>
</tr>
<tr>
<td>PW</td>
<td>Proline Tryptophan</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RSPO1</td>
<td>R-Spondin Homologue 1</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Reverse transcription polymerase chain reaction</td>
</tr>
<tr>
<td>SF1</td>
<td>Steroidogenic factor 1</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
<tr>
<td>SOX9</td>
<td>Sry-related HMG-box 9</td>
</tr>
<tr>
<td>SRY</td>
<td>Sex determining region on Y chromosome</td>
</tr>
<tr>
<td>Stra8</td>
<td>Stimulated by retinoid acid gene 8</td>
</tr>
<tr>
<td>TCS</td>
<td>Testatin/Cres subgroup</td>
</tr>
<tr>
<td>Wnt4</td>
<td>Wingless-type MMTV integration site family, member 4</td>
</tr>
<tr>
<td>WT1</td>
<td>Wilm’s tumor 1 gene</td>
</tr>
<tr>
<td>Wt</td>
<td>Wildtype</td>
</tr>
</tbody>
</table>
INTRODUCTION

Evolution of animal species
The modern theory behind the evolution of species was established by Charles Darwin (1809-1882). He illustrated in *The origin of species* a model for the evolution of species by using a branching tree. Collection and evaluation of paleontological data provide estimates for the minimum constraints on the timing of species divergence. Exact estimates of dates are still difficult to establish due to geological missing links and the sometimes poor quality of the fossil data. However, new fossil findings, reevaluation of old findings together with increased genomic knowledge have constantly improved the estimates of minimum constraints. Benton, Donoghue and Asher (Benton et al. 2008) recently published a newly evaluated summary of paleontological dates for vertebrates (Fig. 1). Throughout this thesis, the dates from this publication are used as references when sexual development in different species is described and results are discussed.

**Figure 1.** Summary of paleontological constraints on vertebrate phylogeny. The original picture was created by Philip C J Donoghue and Michael J Benton, Department of Earth Sciences, Bristol UK (http://www.fossilrecord.net). Dates are minimum constraints in Million years. The mammalian clades, Eutheria, Marsupial and Monotreme and the Eutherian division into Afrotheria, Xenartha and Boreoeutheria are indicated.
Sexual development

The essential purpose of sexual development is to equip the organism with female or male anatomy and physiology to allow for successful reproduction. In vertebrates, sexual development is mainly directed from the gonads. The gonads first develop as sexually undifferentiated tissues with a unique bipotential capacity. At a certain time point sex determination occurs and the gonads develop into either ovaries or testes. Ovaries direct female sexual development and testes direct male sexual development. Sex determination can be initiated by genetic and/or environmental factors. Mammals establish genetic sex at fertilization through heteromorphic sex chromosomes. Most mammals have developed XX for the female sex and XY for the male sex. Sexual development in birds relies on ZZ (male) or ZW (female) sex chromosomes, while reptiles, turtles or fish constitute a nonhomogenous group with respect to sex determining factors. Sex chromosomes among reptiles and fishes consist of both ZW/ZZ and XX/XY systems (Pieau et al. 1999; Volff et al. 2007) and some species are purely temperature dependent while others use both genetic factors and environmental factors in combination.

The development of a testis with similar anatomy and spermatogenesis is well preserved in all vertebrates. However, most birds and some reptilian and monotreme species have evolved a right left asymmetry in their gonad development with the result that only the left ovary develops (Grutzner et al. 2008). In bird, this asymmetry is extended also to the distribution of germ cells with approximately 70% being found in the left gonad of both sexes (Clinton and Haines 1999).

Sexual development in mammals

Mammals are divided into three clades; eutherians (placental mammals), marsupials, and monotremes (Fig 1). Eutherian and marsupial mammals establish chromosomal sex at fertilization via XY or XX sex chromosomes. Monotremes have a unique reproductive system that combines some mammalian (lactation) with some reptilian features (egg laying). They also have a unique sex chromosomal system with 10 sex chromosomes (5X, 5Y), which form a multivalent chain at male meiosis (Grutzner et al. 2004). The monotreme sex chromosomes bear considerable homology to the bird sex chromosomes (Z), but not to the sex chromosomes of marsupial and placental mammals (X), implying that the marsupial and placental sex chromosomes (XY) evolved after the monotreme marsupial species divergence (Rens et al. 2007; Veyrunes et al. 2008).

The bipotential gonads arise in the human embryo at 32 days post ovulation and at embryonic day 9 in mouse embryos, as paired thickenings of the coelomic epithelium on either side of the dorsal aorta. The sex determination gene in humans was identified as a dominant male gene on the Y-chromosome. Extensive chromosomal mapping of 46,XX male individuals revealed a
translocation of a Y-chromosome derived gene. This gene, called SRY is a SOX family transcription factor that was shown to be responsible for complete female to male sex reversal in these individuals, though associated with sterility since some regions on the Y-chromosome are required for spermatogenesis. Moreover, a subset of sex reversed 46,XY females harbored mutations in the SRY gene (Sinclair et al. 1990). Sry has been proven to be the sex determining gene in most eutherian mammals and consistent with this, Sry deficient XY mice show complete male to female sex reversal (Gubbay et al. 1992) and transgenic Sry expression in XX mice results in complete female to male sex reversal (Koopman et al. 1991).

Marsupials also have a Sry gene located to the Y chromosome. However, it has not yet been proven to be the sex determining gene (Foster et al. 1992). The X-derived gene, Atrx, has been found to have a distant relative called Atry located to the Y-chromosomes only in marsupials. Atry has been suggested to be an ancestral sex determining gene in marsupials predating Sry (Pask et al. 2000).

Sex determination in monotremes as platypus still remains a mystery. There is no platypus Sry homologue (Wallis et al. 2007) and the best sex determining candidate gene found so far is DMRT1, which is the candidate sex determining gene in birds. Platypus DMRT1 is localized to the Z end of the X-chromosome chain, which results in double copies in females and a single copy in males (El-Mogharbel et al. 2007). In birds, DMRT1 shows the opposite relation, with two copies in males (ZZ) and a single copy in females (ZW), demonstrating that one can not make a direct connection. Additionally, sex determination at the critical time window is difficult to study in platypus since it is rare, sensitive to environmental changes and a highly protected species.

The phenotypic sex of the embryo is further directed by factors secreted by the differentiating gonads. A dual ductual system with a Müllerian duct and a Wolffian duct is formed in the bipotential embryo (Fig 2). Testosterone produced by the fetal testis is necessary to maintain the Wolffian ducts that along the male pathway develop into male internal genitalia as vas deferens, seminal vesicles and epididymis. Another hormone, Anti Müllerian hormone (AMH), produced and secreted by the testis initiates the regression of the Müllerian ducts in the male embryo that otherwise in the female embryo develop into the female internal genitalia as oviducts, uterus and upper vagina. The Wolffian ducts spontaneously regress in the female embryo that lack fetal production of testosterone (rewieved in Klattig and Englert 2007). Further, testosterone is important for the development of male external genitalia and secondary sex characteristics, while estrogen is important for female external genitalia and secondary sex characteristics. Sex hormones also influence behavioral traits.
Figure 2. Mammalian sexual development begins with the development of a bipotential gonad and a dual ductal system consisting of Mullerian ducts and Wolffian ducts. After sex determination testis development is initiated in XY embryos while ovary development is initiated in XX embryos. Factors secreted by the gonads further direct male or female internal genital development. The Wolffian ducts develop into epididymis, vas deferens and seminal vesicles under the influence of testosterone produced by the testis, while the Mullerian ducts are degraded in the male embryo under influence of AMH secreted by the testis. In the female embryo the Wolffian ducts spontaneously degenerate because of lack of testosterone while the Mullerian ducts develop into oviducts, uterus and upper vagina. This picture shows the shape of the internal genitalia in mice. The uterus in humans is not divided into two parts.
Marsupial and eutherian mammals show some fundamental differences in anatomy of the reproductive organs. Female marsupials lack placenta and have two lateral vaginas with one single birth channel. The marsupial gestation period is very short and the newborn babies are born with the gonads in a sexually undifferentiated stage (Mackay et al. 2004). In general, sexual development in marsupials shows both similarities and differences to eutherians. The timing of sex differentiation and sexual development are different, however Sry might be the sex determining gene, testosterone and AMH are produced by the testis and the Wolffian ducts develop as in eutherians.

**Sex determination in eutherian mammals**

Sex determination is initiated at the time point when Sry is expressed in the mammalian bipotential gonads, which occurs at E10.5 in mice and at E41 in humans (Hanley et al. 2000). Even if the discovery of SRY in 1990 was an important milestone in the field, there was almost a decade until further understanding of the downstream events took place. The first cell type that differentiates after sex determination is the supporting cell lineage, which in the male pathway becomes Sertoli cells and in the female pathway becomes granulosa cells. Initially, it was thought that Sry would bind several target genes in Sertoli cells and trigger the testis program while female development was regarded as the default pathway. Several male specific factors were isolated after the identification of Sry and one of the most important is Sox9, a SOX family transcription factor, related to Sry. Sox9 is heavily upregulated in Sertoli cells after the peak expression of Sry, which was shown to be crucial for Sertoli cell differentiation. It has also been shown that upregulation of Sox9 without Sry is sufficient to trigger the male pathway (Chaboissier et al. 2004) and 46,XX individuals with SOX9 duplications show female to male sex reversal (Huang et al. 1999). Conversely 46, XY individuals with haploinsufficiency of SOX9 develop campomelic dysplasia, a skeletal disorder associated with male to female sex reversal (Schafer et al. 1995). Further, Fgf9 knockout mice were reported to show partial to complete male to female sex reversal which suggested that Fgf9 also plays an important role in testis formation (Colvin et al. 2001). Basically no candidate genes for female sex determination were found during this period.

However, the last years have seen tremendous progress in that regard. At the same time as several female upregulated factors were identified a new exciting picture of sex determination emerged. It is now clear that the female and male pathways are linked through the Wnt canonical signaling pathway and female and male factors act in an antagonistic fashion to tip the balance into either the female or the male differentiation program (Fig 3) (Capel 2006). The signaling molecule Wnt4 and the transcription factor Foxl2 were the first female upregulated and male downregulated genes identified, and both proved to be necessary for proper ovary differentiation. Wnt4 female knockout mice show male specific characteristics, such as the
presence of steroid producing cells and development of a male vascular system (Jeays-Ward et al. 2003; Vainio et al. 1999). Foxl2 is required for normal female fertility and patients with mutated FOXL2 have the BPES syndrome including premature ovarian failure (POF) and eyelid abnormalities (Crisponi et al. 2001). Female Foxl2 knockout mice are sterile and show partial testis formation mainly due to upregulation of male factors such as Sox9 in the female supporting cells (granulosa cells) (Ottolenghi et al. 2005). Although neither Wnt4 or Foxl2 cause complete XX female to male sex reversal when deleted, a study with a Wnt4 and Foxl2 double knockout mouse showed almost complete XX female to male sex reversal (Ottolenghi et al. 2007). Therefore the discovery of a 46, XX individual with complete female to male sex reversal associated with palmoplantar hyperkeratosis and a recessive mutation in the gene R-spondin 1 (RSPO1) (Parma et al. 2006) was very important. It became clear that also one single female gene can tip the balance towards the female pathway. Rspo1 is up regulated in the ovary at the critical stage of sex determination and generation of Rspo1 knockout mice proved that Rspo1 is on the top of the female pathway (Chassot et al. 2008).

**Figure 3.** β-catenin signaling in the supporting cell type. Female RSPO1 stabilizes β-catenin through the Wnt signaling pathway (Capel 2006), while Sox9 in the male pathway has been shown to promote β-catenin destruction (Akiyama et al. 2004). As a transcription factor, Sox9 has also been shown to upregulate Fgf9 to further promote the male pathway and the signaling molecules Fgf9 and Wnt4 are suggested to act as negatively regulators of each other, suppressing the opposite pathways (Kim et al. 2006).
Mouse testis formation

There are two basic functions for the testis: one is to provide an environment for the maturation of germ cells and the second is to produce and export hormones that further masculinize the embryo. There are five major cell types in the fetal testis namely germ cells, Sertoli cells, peritubular myoid cells, Leydig cells and vascular endothelial cells (Fig 3).

Germ cells are the progenitor cells of adult spermatozoa, which ensure transmission of genetic information from one generation to the next. Embryonic germ cells migrate through the hindgut and enter the gonad already before sex determination. After sex determination, the first visible sign of testis differentiation is the formation of testicular cords, where Sertoli cells migrate to enclose, protect and nurse the germ cells in a cord like structure (Fig. 4). A basal membrane consisting mainly of collagen, laminin and proteoglycans is produced by Sertoli cells and peritubular myoid cells (Skinner et al. 1985), to further establish the cords (Fig 4).

The first difference between male and female germ cell behaviour is their timing of onset into meiosis. Male embryonic germ cells enter mitotic arrest at around E13.5 in mouse while female germ cells enter meiosis at the same time. It was recently shown that Retinoic acid (RA) is responsible for induction of meiosis through stimulation of the gene \( Str8 \) in female germ cells. RA is produced by the adjacent mesonephros (Fig 4) and diffuses into the gonads. In the male gonads RA is locally degraded by production of the enzyme Cyp26b1 secreted by Sertoli cells (reviewed in Swain 2006).

Testis differentiation, in contrast to ovary differentiation, is characterized by strong cell proliferation and cell migration from the adjacent mesonephros. Endothelial cells enter the gonad from the mesonephros to form a male arterial network (coelomic vessel) important for transport of hormones produced in the testis (Fig 4) (Brennan et al. 2002). The space between the cords is called interstitium and is filled with steroid producing Leydig cells and vascular endothelial cells.

As described earlier, Sertoli cells play a central role in initiating the male pathway. They continue to direct testis formation by secreting many of the factors necessary for cord formation and cell migration from mesonephros. Most evidence also suggests that Sertoli cells induce differentiation of other testis cell types such as peritubular myoid cells and fetal Leydig cells (Clark et al. 2000; Pierucci-Alves et al. 2001).
Figure 4. Testis formation. Sertoli cells play a central role in testis formation. They migrate to form the characteristic testicular cords, to nurse and protect the germ cells. They also secrete several of the important factors needed for cell migration from the adjacent mesonephros and differentiation of other cell types, as peritubular myoid cells that enclose the cords and fetal Leydig cells that produce testosterone. The male arterial system is important for transport of hormones to the rest of the embryo.

Molecular interactions in mouse testis differentiation
Since the discovery of Sry, several transcription factors have been identified downstream or upstream of Sry including Sf-1, Wt1 +/-KTS isoforms, Sox9, GATA4, Pod1 and the nuclear receptor protein Dax1. Analyses of genetically engineered mice lacking these factors have been very important in order to understand the underlying mechanisms that direct testis formation. Sf-1 and Wt1-KTS knock-out mice show regressed gonads of both sexes (Hammes et al. 2001; Luo et al. 1994), demonstrating that Sf-1 and Wt1-KTS are both required for the development of the bipotential gonad. Transcriptional activation of Sry is crucial for initiation of testis formation and in vitro studies have shown that SF-1 can activate transcription of Sry (Pilon et
Further, Wt1-KTS deficient mice showed XY male to female sex reversal due to severely reduced levels of Sry (Hamnes et al. 2001). Other transcription factors have been suggested to be important for activation of Sry as well. However, genomic evolutionary studies imply that factors that mediate Sry transcription might differ to some extent between different mammalian species (Ito et al. 2005). As mentioned earlier, upregulation of Sox9 immediately after Sry is necessary to establish the male pathway. Until recently the link between Sry and Sox9 was unclear. Interestingly, a recent study shows that SF-1 and Sry in combination are sufficient to activate transcription from a highly conserved sequence element in the Sox9 promoter (Sekido and Lovell-Badge 2008).

Induction of AMH expression is important for Müllerian duct regression. SF-1 together with GATA4 and WT-1 has been shown to induce expression of AMH in Sertoli cells (Nachtegal et al. 1998; Shen et al. 1994; Viger et al. 1998). SF-1 is a multifunctional transcription factor during testis differentiation and in addition to the described functions above SF-1 is also involved in regulating steroid biosynthetic genes in fetal Leydig cells (Reinhart et al. 1999). DAX1 was first identified as a female gene due to the fact that duplication of the DAX1 locus in XY individuals cause complete male to female sex reversal (Bardoni et al. 1994). Overexpression of Dax1 in XY transgenic mice with a weak allele of Sry results in the same phenotype (Swain et al. 1998). Further experiments have also shown that Dax1 is not required for normal ovary development. On the contrary Dax1 was shown to be required for normal testis differentiation in a dosage sensitive manner (Bouma et al. 2005).

Pod1 is a transcription factor thought to repress the expression of SF-1. Overexpression of SF-1 in Pod1 knockout mice results in a dramatic overexpansion of fetal Leydig cells with improper steroid production (Cui et al. 2004).

Signal transduction between cells through a number of well conserved pathways is fundamental for cell differentiation during embryogenesis. Around the time of discovery of Sry and Sox9, the main research focus was on transcription factors. More recently, signal transduction pathways became the focus of study. So far, five signaling pathways have been shown to play a role in testis formation, including Fgf9, the insulin signaling family (Ir, Igf1r and Irr), Dhh, Pdgfr-α and Notch.

Proliferation of Sox9 positive Sertoli cells is crucial to override female signals and restore the male pathway. Fgf9 ablation results in partial to complete male to female sex reversal in knockout mice (Colvin et al. 2001) and Fgf9 is suggested to signal through Fgfr2 and promote Sertoli cell proliferation (Schmahl et al. 2004). A recent study shows that conditional knockout
mice for Fgfr2 display an early disruption of Sertoli cell proliferation causing male to female sex reversal (Kim et al. 2007).

The role of insulin signaling in fetal testis formation was shown in XY mice mutant for all three receptors (Ir, Igf1r and Ir) that developed complete male to female sex reversal (Nef et al. 2003). Dhh is secreted by Sertoli cells and acts via the Ptc receptor that is expressed by peritubular myoid cells (Clark et al. 2000; Pierucci-Alves et al. 2001). Dhh null XY mice are sterile and show different degrees of pseudohermaphroditism depending on genetic background. A closer examination of the fetal testis of Dhh null mice revealed impaired deposition of the basal lamina between Sertoli and peritubular myoid cells (Pierucci-Alves et al. 2001).

Characterization of Pdgf knockout phenotypes in fetal testis identified Pdgfr-α as a critical mediator of signaling in the early testis at multiple steps. XY gonads displayed disruptions in organization of testicular cords and vasculature as well as reduction in cell migration, cell proliferation and fetal Leydig cell development (Brennan et al. 2003).

Gonad culture systems where Notch signaling was either blocked or overstimulated strongly suggest that Notch signaling is involved in fetal Leydig cell renewal in fetal testis. It also showed that regulation of the fetal Leydig cell population influences germ cell survival and testis cord formation (Tang et al. 2008).

**Disorders of sexual development (DSD) in XY individuals**

Abnormalities in sexual development in humans are collectively referred to as DSD. DSD can be divided into different forms, largely depending on the underlying genetic abnormalities (Lee et al. 2006). Since testatin is expressed in fetal mouse testis, the focus in this section will be on the group with disorders of testicular development (46,XY gonadal DSD), in particular complete gonadal dysgenesis. This is a rare group of phenotypically female individuals who are often diagnosed due to primary amenorrhea and lack of pubertal development. Further examinations show a 46,XY karyotype, high FSH levels and undifferentiated (streak) gonadal tissue. The undifferentiated gonads fail to produce AMH and testosterone in the fetal stage that would have been necessary for development of male internal and external genitalia and for regression of Müllerian structures. Therefore, female internal and external genitalia are formed. Complete 46, XY gonadal dysgenesis is associated with a risk of developing germ cell tumors and gonadectomy is therefore recommended in these individuals (Looijenga et al. 2007).

The genetic background of gonadal DSD in XY females is only found in 20-50% of the cases. Around 10-15% are due to mutations or deletions of the SRY gene (reviewed in Val and Swain...
Duplications of the \textit{DAX1} gene on Xp21.2 seem to account for a similar proportion of cases (Barbaro et al. 2007 and unpublished data). Rarer causes are mutations or deletions of \textit{SF1}, \textit{WT1}, \textit{DHH}, \textit{SOX9}, \textit{ARX}, and the \textit{DMRT} genes on distal 9p. Most of these latter genetic aberrations are associated with symptoms from additional organ systems and cause more complex syndromes including gonadal dysgenesis. Many genes that have been studied in mouse models are also found to cause gonadal DSD in humans (Val and Swain 2005). However species related differences exist, especially in the timing and level of expression of genes such as \textit{Sry/SRY}, \textit{Sox9/SOX9} and \textit{Dax1/DAX1} (Hanley et al. 2000). \textit{Fgf9} is an interesting example that has been shown to play a very important role in mouse testis differentiation but despite the fact that patients with mutations in \textit{FGF9} have been described there is no report on gonadal DSD in these patients (reviewed in Val and Swain 2005) indicating species differences in this pathway.

Sex chromosome mosaicism such as 45,X0/46,XY, which may be present in gonads but remain undetected in blood, is likely to underlie a proportion of unexplained cases with 46,XY gonadal dysgenesis (Barbaro 2008). However, it is clear that a substantial proportion of all patients cannot be explained by known genetic defects. It is thus likely that additional genes remain to be identified, which are specifically involved in early testis differentiation and which cause abnormal testis development when mutated.

\textbf{Testatin}

Finding new genes in sexual development could improve diagnostics for individuals with gonadal DSD, as well as clarify basic mechanisms behind organogenesis. \textit{Testatin} was previously isolated by our group as a novel candidate gene for involvement in early sexual development. A screen was designed to specifically pick up putative extracellular genes, differentially expressed in fetal testis or ovary just after sex determination (Nordqvist and Tohonen 1997; Tohonen et al. 1998). \textit{Testatin} mRNA is strongly upregulated in fetal testis just after the peak expression of \textit{Sry} and at the same time it is downregulated in the female ovary. \textit{In situ} experiments showed specific expression in Sertoli and germ cells in fetal testis and RT-PCR showed continuous expression in adult testis (Tohonen et al. 1998; Kanno et al. 1999). The \textit{testatin} gene codes for a putative extracellular protein of 137 aa with a signal peptid of 31 aa (predicted by Ensembl) and two potential glycosylation sites. The nucleotide sequence shows close homology to the family 2 cystatins within the cystatin super family of protease inhibitors.
Cystatin super family
The cystatin superfamily comprises a large group of proteins with a common cystatin domain structure, which also suggests a common evolutionary origin. Classifications are made largely based on sequence homology, certain structural features such as the presence or absence of disulphide bonds, number of cystatin domains or functionality. Originally, the family was divided into three classes. These classes are family 1 (stefins) that are small intracellular proteins that lack conserved cysteins, family 2 (cystatins) that are small secreted proteins with conserved cysteins and family 3 (kininogens) that are large secreted multi functional proteins with conserved cysteins. All three original classes show strong conservation in three consensus sites and are proven tight reversible inhibitors of cystein proteases such as papain and cathepsin B, H and L in vitro (Barett 1986). However, identification of other related cystatin domain containing proteins such as fetuin and the histidine-rich glycoprotein, that lack conservation in the three consensus sites and most likely are not cystein protease inhibitors, have lately increased the number of classes in the cystatin super family (Brown and Dziegielewska 1997). Members of the cystatin superfamily are implicated in a number of physiological mechanisms as antigen presentation, bone remodeling and cancer (reviewed in Turk et al. 2008).

Family 2 cystatins
Family 2 cystatins are 13-15 kDa proteins with approximately 120 aa and a putative signal peptide of 20-26 aa. Seven human family 2 cystatins have been identified namely CST1/cystatin SN, CST2/Cystatin SA, CST3/Cystatin C, CST4/Cystatin S, CST5/Cystatin D, CST6/Cystatin E/M and CST7/Cystatin F (reviewed in Turk et al. 2008). They are all proven tight reversible inhibitors of cystein proteases such as papain and cathepsin B, H and L in vitro (Barett 1986) and show high conservation in the three consensus sites. Solving the crystal structure of Cystatin C (Bode et al. 1988; Janowski et al. 2001) was instrumental for understanding its mechanism of action. A wedge shaped structure consisting of three highly conserved consensus sites, a G residue in the N-terminal part, a QXVXG site in the central part and a PW site in the C-terminal part of the protein binds and blocks the active site of the protease (Stubbs et al. 1990).

The family 2 cystatins have due to a broad expression and secretion in a variety of different organs and body fluids (Abrahamson et al. 1986; Dickinson et al. 2002; Jiborn et al. 2004) been implicated in different general physiological processes. Cruzipan, a protease produced by a protozoic parasite (Trypanosoma cruzi) is strongly inhibited by human Cystatin C in vitro, suggesting a role in defence against infection (Stoka et al. 1995). Extracellular cystatins have also been suggested to protect against tissue damage induced by released lysosomal cathepsins (Kuester et al. 2008; Travis 1988). Altered expression levels of cystein protease inhibitors
relative to cathepsins in cancerous tissues may permit accelerated extracellular matrix
destruction and tumor invasion (Kos et al. 2000; Sokol and Schiemann 2004).

Wildtype (wt) cystatin proteins are in general prone to form amyloid fibrils and a single
mutation in human Cystatin C (L68Q) causes massive amyloid depositions in the arteries of
patients with the autosomal dominant disorder Heriditary Cystatin C Amyloid Angiopathy
(HCCAA) (Levy et al. 1989). The outcome of HCCAA is fatal due to multiple strokes. Several
mechanisms have been suggested to be involved in amyloid formation of cystatins. One is
dimerization via three-dimensional domain swapping, which was demonstrated to occur in wt
crystallized human cystatin C and was greatly facilitated in the L68Q variant (Janowski et al.
2001). However, experiments that prevented dimerization via three-dimensional domain
swapping in the L68Q variant showed that 20% of the protein still formed amyloid fibrils
suggesting that also other mechanisms might be involved in amyloid formation (Nilsson et al.
2004). Fibril formation of the family 1 protein stefin B in vitro was shown not to be compatible
with the three-dimensional domain swapping model. Instead a model called “hand shaking”
was proposed to promote fibril formation with trans to cis isomerization of Pro74, that is widely
conserved throughout the cystatins (Jenko Kokalj et al. 2007).

A polymorpism in the signal peptide of Cystatin C has been associated with increased incidence
of Alzheimers disease (Cathcart et al. 2005; Chuo et al. 2007). However, the mechanism behind
this association is probably not, as one first would have assumed, linked to its ability to form
amyloid fibrils. Instead, this association is suggested to be due to a protective role of Cystatin C
in formation of amyloid-β aggregates. Cystatin C has been reported to associate with amyloid-β
in vitro and a recent study showed that Cystatin C binds amyloid-β in cerebrospinal fluid, blood
and brain and inhibits its formation into insoluble aggregates in vivo in mice (Mi et al. 2007). At
the same time the polymorphism in the signal peptide is associated with reduced levels of
Cystatin C in serum (Chuo et al. 2007).

Testatin/Cres subgroup of family 2 cystatins
Cres was isolated as the first Testatin/Cres subgroup gene in search for novel fertility genes
expressed in epididymis (Cornwall et al. 1992). A few years later Testatin was isolated as a
candidate gene for involvement in early sexual development (Tohonen et al. 1998). Thereafter,
several similar genes with reproductive tissue specific expression and close sequence homology
to the family 2 cystatins were reported and it became clear that a new reproductive tissue
specific subgroup within family 2 cystatins had been identified (Cornwall and Hsia 2003). The
Testatin/Cres subgroup genes show a rather low sequence identity ranging from 17-45% (Paper
II). However, common structural features as lack of conservation in the N-terminal and central
consensus motives but retained conservation in the C-terminal PW-site, as well as in the four C-
residues that are conserved among family 2 cystatins, suggest a common origin but different specificity compared to the classical family 2 cystatins. Indeed, Cres was shown to selectively inhibit the serine protease prohormone convertase 2 in vitro, instead of the classical cystatin targets papain and cathepsin B (Cornwall et al. 2003). The reproductive tissue specific expression of the Testatin/Cres subgroup genes, which is in striking contrast to the broadly expressed family 2 cystatins, together with a different functional specificity strongly suggest that the Testatin/Cres subgroup genes have evolved to perform specialized functions within the reproductive system.

Today, eight mouse genes in the Testatin/Cres subgroup have been identified (Fig. 5). Expression in adult testis and/or epididymis implicates specialized functions during spermatogenesis and/or sperm maturation. Cres/Cst8 protein is localized to spermatids, spermatozoa and is secreted by epithelial cells in the proximal region of epididymis (Cornwall and Hann 1995; Cornwall et al. 1992; Syntin and Cornwall 1999; Yuan et al. 2007). CstTE-1/Cres3 protein is localized to the cytoplasm of Sertoli cells in testis and epithelial cells in the proximal region of epididymis (Li et al. 2005). Different studies have shown mRNA expression of CstT/Cst13 in pachytene spermatocytes and round spermatids (Shoemaker et al. 2000), CstE2 in caput epididymis (Li et al. 2003), CstE1/Cres2 in epididymis (Hsia and Cornwall 2003), CstSC in Sertoli cells in the testis (Li et al. 2002), CstL1/Rcetv1/Rcetv2 in spermatagonia, spermatocytes, round spermatids and epididymis (Xiang et al. 2008) and Testatin/Cst9 in Sertoli and germ cells in the testis (Kanno et al. 1999; Tohonen et al. 1998).
Figure 5. Eight mouse Testatin/cres genes are identified, all with a reproductive tissue specific expression in contrast to the classical family 2 cystatins. Phylogenetic relations of the individual eight mouse genes are here shown in a tree together with Cystatin C.

In light of the above, it seems plausible that the Testatin/Cres subgroup of genes has evolved to exert specific but overlapping functions within the reproductive system. Ancestors of the cystatin super family proteins can be traced back to plants where proteins related to stefins are found (Brown and Dziegielewska 1997; Muller-Esterl et al. 1985; Rawlings and Barrett 1990). Family 2 cystatins are reported in different vertebrate species such as chicken and mammals. Regarding the Testatin/Cres subgroup of genes, primate members have been reported previously such as human and Macaca mulatta CST11 protein that is localized to the lumen throughout epididymis (Hamil et al. 2002), human CRES/CST8 protein that is localized to sperm equatorial segment (Wassler et al. 2002) and two human testatin orthologue genes. One of these, CSTL9, was isolated by our group and showed testis specific mRNA expression (Eriksson et al. 2002) while the other CST9/CLM, was isolated from bone marrow stromal cells (BMSC) and reported to be expressed in several human tissues such as heart, placenta and lung, liver, skeletal muscle and pancreas (Sun et al. 2003).
AIMS

The specific aims of this thesis were

- To study the *in vivo* role of testatin in early sexual development by generating and analyzing a *testatin* knockout mouse. Testis differentiation, male fertility, body weight and hormonal parameters were analyzed in *testatin* deficient mice (Paper I and unpublished project A).

- To investigate the possibility of functional redundancy between the Testatin/Cres subgroup genes. Spatial and temporal mRNA expression patterns were studied for the Testatin/Cres subgroup genes, in normal as well as in *testatin* knockout mice (Paper II).

- To evaluate the mRNA expression patterns of all human Testatin/Cres subgroup genes and to study their possible involvement in complete 46,XY gonad dysgenesis (Paper III and unpublished project B).

- To perform an evolutionary study of the Testatin/Cres subgroup genes in order to understand when and how the Testatin/Cres subgroup genes emerged (Paper III).

- To produce recombinant testatin protein that in future projects can be used for antibody production, crystallization studies and protease inhibitor studies (unpublished project C).
METHODS AND MATERIALS

This section summarises the methods included in this thesis. I started to work in the team after the Testatin knockout mice had been generated. My part in this work started with the analysis of the Testatin knockout mice. For the parts of my work that have resulted in publications, all methods are described in detail in each paper. However, some results are unpublished and the methods used in those studies are described in this section. In total, the following methods were used in the work included in my thesis: Generation of a testatin knockout mouse\(^1\), Mouse breeding and genotyping\(^{1,IIA}\), RT-PCR\(^{I,III}\), In situ hybridization\(^{II}\), Histology\(^{II}\), Immunohistochemistry\(^I\), Electron microscopy\(^I\), Fertility assessment in mice\(^I\), Preparation and analysis of mouse epididymal spermatozoa\(^I\), Hormone measurements\(^I\) (testosterone, LH and FSH), Fetal and adult gonad collection\(^{I,II}\), Backcross of testatin knockout mice\(^A\), Body weight assessments\(^A\), RNA extraction\(^II\), Real-time PCR\(^II\), Cloning\(^C\), Eukaryotic protein expression\(^C\), DNA sequencing\(^{II,III}\), and Bioinformatics\(^{I,II,III}\). The superscript numbers indicate the publication were the protocols are described. Previously unpublished methods are indicated with superscript letters.

**Unpublished project A, body weight development in testatin deficient mice**

Mouse breedings and genotyping were performed as described earlier in paper I and II. Body weight measurements were performed regularly. For body weight study part 1, consisting of animals on a mixed genetic background, three different groups of six male mice were weighed at ages 3, 10.5 and 15 months. For body weight study part 2, which consisted of 90% backcrossed (generation five) animals, 10 wt male mice and 10 Testatin knockout mice were sent to Gothenburg (Swegene Center for mouse Physiology and bio-imaging) for regular weight and DEXA analyses. This group of animals were weighed every 3 months until 6 months of age and thereafter every month until 15 months of age. In body weight study part 3, consisting of 100% backcrossed animals (generation 10), both male and female mice were weighed every week from three weeks of age until 4-6 month of age.

**Unpublished project A, backcrossing of the mice**

To create fully backcrossed testatin knockout animals, male heterozygous (+/−) Testatin mice were bred 10 generations with C57BL6 female mice. Each generation of male mice were genotyped (as in paper I and II) and heterozygous males were selected for subsequent breeding with wildtype C57BL6 females. The Y-chromosome of C57BL6 was introduced at generation five, when heterozygous females were bred with C57BL6 males.
Unpublished project B, screening for mutations in the Testatin/Cres subgroup genes in disorders of sexual development

Genomic DNA from eight patients with complete 46,XY gonadal dysgenesis (XY females) was sequenced. The following genes were sequenced (ensembl 2008): CST11/TCS2 (ENSG00000125831), CRES/TCS5 (ENSG00000125815), CST9/TCS7A (ENSG00000173335), CST9L/TCS7B (ENSG00000101435), all expressing functional transcripts as determined in Paper III. PCR reactions using 50 ng DNA for each exon was set up using EXT DNA polymerase (Finnzymes), 0.3 mM primers and 200 µM of each dNTP in a total volume of 25 µl. For primer sequences and annealing temperatures for each gene and exon see Table 1. Two to five µl PCR reaction was cleaned prior to sequencing (depending on the intensity of the fragment on agarose gel), in a total volume of 6,5 µl together with Exonuclease I (0.5 µl) (Promega) and Shrimp alkaline phosphatase (1 µl) (Promega) at 37 °C for 15 min, followed by 15 min heat inactivation at 85 °C. Sequencing was performed using the ABI Big Dye Termination v3.kit and the 3830 DNA Analyzer (Applied Biosystems).

Table 1. Primers used for sequencing patients with gonadal DSD.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Promotor /Exon</th>
<th>Forward primer 5´-3´</th>
<th>Reverse primer 5´-3´</th>
<th>Temp</th>
</tr>
</thead>
<tbody>
<tr>
<td>CST9L/TCS7B</td>
<td>Pr.</td>
<td>GCTGGCAAAATAGCTCACC</td>
<td>GTGGAGTCATCTCATCTGCC</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>Ex. 1</td>
<td>CGAGATGTCCCCAGGGTCTTA</td>
<td>TTTCTGTCAGCCCTTTTCC</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>Ex. 2</td>
<td>GGAGAATGTTGGGACTTCTC</td>
<td>TCTACTCTCTAAGTCTC</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>Ex. 3</td>
<td>GCAGCACATGTAGCCAGAG</td>
<td>TGACTACTTGTGCCACCTG</td>
<td>58</td>
</tr>
<tr>
<td>CST9/TCS7A</td>
<td>Ex. 1</td>
<td>GACAGGGTAACTGCAGCGA</td>
<td>CTGCTGTGGTAGAACCAGT</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>Ex. 2a</td>
<td>CATGGGCAAGCTCTCAATATG</td>
<td>GTACCACAGACATTGTCAC</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>Ex. 2b</td>
<td>CATTCTCTCTCTCTCTGTG</td>
<td>GACACAGTCAAATAGCC</td>
<td>55</td>
</tr>
<tr>
<td>CST11/TCS2</td>
<td>Pr.</td>
<td>GTCCAACCAGAAAGCTCCTGT</td>
<td>GTGCTCACATCAAGCTAAGC</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>Ex. 1</td>
<td>GTCCAACCAGAAAGCTCCTGT</td>
<td>GTGCTCACATCAAGCTAAGC</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>Ex. 2</td>
<td>CCTAGGACAGCTCCTGTC</td>
<td>TAGGAAGGGATGTTACCC</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>Ex. 3</td>
<td>GCCCTTCTCCATCTCCTTG</td>
<td>GTAGCTTACCTTTGGAGC</td>
<td>58</td>
</tr>
<tr>
<td>CRES/TCS5</td>
<td>Pr.</td>
<td>GCCCCAGAAGGCTCTCTG</td>
<td>CAGTGGTCTCCTCAGGTGTC</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>Ex. 1</td>
<td>CATCACAAGGGTAAACCTTC</td>
<td>CCAGTAAAGGAGTCTCCT</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>Ex. 2</td>
<td>TCCATGGGAAAGGAGGAC</td>
<td>GCAATGTCACCTTATAGG</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>Ex. 3</td>
<td>TTGTTTGACACAGGCTCTCA</td>
<td>GGGCCTGGATCTCTGCTAT</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>Ex. 4</td>
<td>CTAAGCTCTGCAGATTCTCCTG</td>
<td>ACCA TCAAGCTGAAACAGG</td>
<td>58</td>
</tr>
</tbody>
</table>

The annealing temperature used in PCR is indicated with Temp.
Unpublished project C, production of recombinant protein in eukaryotic cells

Mouse Testatin and Cystatin C cDNA was cloned to pCEP-Pu/AC7 (generated and received with permission from Takako Sasaki, Timple lab Munich, Germany). cDNA was picked up with RT-PCR as described in paper III from RNA isolated from mouse testis, as described in paper II. The following primers were used: Mouse Testatin forward primer, 5’CGACTTAAGCTTGCCTCCATGTCCTGTCCACTG 3’, reverse primer, 5’CGACTTCTCAGCTAGTGATGGTGATGGTGATGAGCAACACACTCTGAAGAGCTTTG and mouse Cystatin C forward primer, 5’CGACTTAAGCTTCCACCATGGCCAGCCCGCTG 3’and reverse primer, 5’GGACTTCTCAGCTAGTGATGGTGATGGTGATGGCATTTTTGCAGCTGAATTTTGTC. A his-tag in the N-terminal end of the cDNA was introduced with the reverse primer. The natural signal peptide of both Testatin and Cystatin C was kept in the cDNA to allow for secretion after expression. The natural Kozak sequences were kept as well. The PCR products were cloned with XhoI and HindIII (New England Biolabs) to the pCEP-Pu/AC7 vector. Correct constructs were sequenced using the ABI Big Dye Termination v3.1 kit and the 3830 DNA Analyzer (Applied Biosystems) and transfected to EBNA 293 cells (ATCC number CRL-10852). Cells were grown in DMEM Glutamax (Invitrogen) together with 10% fetal calf serum (Invitrogen) and 0.25 mg/ml G-418 (Invitrogen). Confluent cells were trypsinized with Dulbecco’s PBS without Mg²⁺ and Ca²⁺ (Invitrogen) and trypsin EDTA (Invitrogen). 60-80% confluent cells were transfected with 1-2 mg sterile precipitated vector DNA using FuGENE6 (Roche) according to manufacturer’s protocol. Transfected cells were grown in medium plus puromycin 2 mg/ml as selection for the vector. Expressed protein was isolated from the cell supernatant with nickel charged resin NTA agarose beads (Qiagen) according to the manufacturer’s protocol. Protein was investigated by Western blot using Novex gel system, NuPAGE, Bis tris gels 12% MOPS buffer (Invitrogen) and hybond P PVDF protein binding membranes (Amersham) for blotting. Blotted membranes were blocked with TBS (5mM Tris and 20 mM NaCl PH 7.6) and 10% milk powder. Detection was performed with monoclonal anti-poly histidine antibodies (R&D systems) and a horseradish peroxidase-conjugated secondary antibody (Santa Cruz Biotechnology).

Unpublished project C, production of recombinant protein in prokaryotic cells

Human Testatin/CST9L/TCS7B cDNA (ENSG00000101435) was isolated from human testis RNA (Clontech) with reverse transcriptase PCR as described in paper III with forward primer 5’CGACTTGCTAGCGCCTGGCATTTCCACGAGCAAGGG 3’ and reverse primer 5’GGACTTCTCAGTGAGCTTTCCACGAGCAAGGG 3’, cleaved with NheI and XhoI (New England Biolabs) and cloned to vector pET-28a-c(+) (Novagen). The natural stop codon was included which allowed a his-tag only in the N-terminal end. Mouse Testatin was amplified
from a vector containing the mouse cDNA sequence (ENSMUSG00000027445) with the following primers; forward primer 5´CGACTTGCTAGCGCCAATAAAGAAAACAAACAGATCTG 3´ and reverse primer 5´GGACTTCTCGAGCTATGACACACACTCCTGAAGAGC 3´, thereafter cleaved and cloned to vector pET-28a-c(+) (Novagen) as described with human Testatin. These clones were handed over to prof. Ylva Lindqvist and collaborators (Jodie Guy and Maria Svärd) at MBB, KI for expression in the cytoplasm of several different E.coli strains. Very little soluble protein was produced in all E.coli systems that were used. Therefore a different approach for periplasmic expression was initiated. Mouse and human Testatin were isolated from the previously made clones with new primers; mouse Testatin forward primer 5´CGACTTCCATGGATGCCAATAAAGAAACAAACAGAT 3´ and reverse primer 5´GGACTTCTCGAGTGAGCACACACTCCTGAAGAGC 3´ and human Testatin forward primer 5´CGACTTCCATGGATGCCTGGCATTTCCACGAGCAAAGGG 3´ and reverse primer 5´GGACTTCTCGAGTGGAATCCCTCCAAGCAGG 3´. PCR products were cleaved with NheI and XhoI (New England Bioloabs) and cloned to vector pET-22b(+) (Novagen). pET22 contains a pelB leader in the 5´ end, that allows for transport to the periplasm of E.coli. pET22 also contains a his-tag in the 3´end, that is expressed in constructs without the natural stop codon in the cDNA sequence. We also made constructs with the natural stop codon to have the possibility to express protein without his-tag as well. Clones were handed over to prof. Ylva Lindqvist and collaborators (Jodie Guy and Maria Svärd) at MBB, KI for further expression in different E.coli strains. Expression is currently under evaluation.
RESULTS

Normal sexual development and fertility in testatin knockout mice (Paper I)

A testatin knockout mouse was generated and analyzed to evaluate the role of testatin in early testis formation. We found no abnormalities in the testatin knockout mice with regard to fetal testis morphology or cellular ultra structure. Testis cords were properly formed and all the celltypes; Sertoli cells, germ cells, Leydig cells, peritubular myoid cells and the basal lamina were properly differentiated both in E17.5 (Fig. 3, 4 and 5. Paper I) and in E 15.5 testis (unpublished results).

Adult male mice showed normal body, testis and epididymis (cauda) weight, number of offspring, spermatogenesis (Table 2. Paper I) and hormonal parameters (testosterone, LH and FSH) (Table 3. Paper I).

Body weight in testatin deficient mice (unpublished project A)

We initially observed a lower body weight in the testatin deficient (td) mice analyzed in publication I, therefore an extended body weight study in three parts was performed.

The first part was performed with animals on a mixed genetic background (the same animals as in publication I). Body weight was measured in td and wt control male mice at different ages. No significant difference in body weight was found. However, a remarkable but not significant lower body weight and a higher testosterone level was observed in 15 month old td male mice compared to wt controls (Table 2 and 3). On the other hand, 15 month old td and wt mice did not show any significant difference in testis and cauda weight, number of offspring, sperm count or sperm motility. A possible explanation was that td mice gained less fat with age compared to wt controls due to hormonal differences. Therefore a second body weight study was initiated.

The testatin knockout strain was backcrossed to C57BL6 and generation 5 (approximately 90% backcrossed) was analyzed in study 2, performed in collaboration with Swegene Center for mouse Physiology and bio-imaging in Gothenburg. Body weight and fat percentage (analysed by DEXA) were measured regularly in 10 td and 10 wt controls from age 3 to 16 months. Td mice showed a significantly lower body weight of 7- 9% from 3 month until 8 month of age compared to wt controls (Table 4). Percentage of body fat was however not significantly different between the groups (Table 5). From 9 months to 16 months there was no significant body weight difference anymore however there was a similar difference observed as in part 1. Therefore, we concluded that td mice may have a reduced body weight already as young mice,
but since body weight probably depends on many different factors a strong power is required to
detect this difference and we may have missed this difference in study 1 due to lower power
(n=6 on a mixed genetic background in study 1 compared to n=10 on 90% backcrossed animals
in study 2).

A third body weight study was conducted to investigate if the difference seen in study 2 was
true. Fully backcrossed wt and td mice (n= 8-13), both female and male animals were weighed
regularly. This time we could not see any difference in body weight between the wt and td
groups in either female or male mice (Table 6). Therefore, we conclude that there is no obvious
body weight difference between td mice and wt mice.

**Table 2. Body weight in animals with mixed genetic background, study A part 1**

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Body weight (g) at Age (months)</th>
<th>3m</th>
<th>10.5m</th>
<th>15m</th>
</tr>
</thead>
<tbody>
<tr>
<td>Testatin -/-</td>
<td></td>
<td>33.3 ± 0.803</td>
<td>39.17 ± 2.822</td>
<td>41.34 ± 2.376</td>
</tr>
<tr>
<td>Testatin +/-</td>
<td></td>
<td>31.2 ± 1.249</td>
<td>36.17 ± 2.915</td>
<td>48.17 ± 2.330</td>
</tr>
</tbody>
</table>

All values are means ± standard errors of means, n = 6 for all groups, p > 0.05 for all values (t-test).

**Table 3. Testosterone levels in adult testis in animals with mixed genetic background, study A part 1**

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Intratesticular testosterone (pg / 100 µl) at Age (months)</th>
<th>3m</th>
<th>15m</th>
</tr>
</thead>
<tbody>
<tr>
<td>Testatin -/-</td>
<td></td>
<td>1575 ± 916</td>
<td>1396 ± 735</td>
</tr>
<tr>
<td>Testatin +/-</td>
<td></td>
<td>1887 ± 885</td>
<td>449 ± 102</td>
</tr>
</tbody>
</table>

All values are means ± standard errors of means, n = 6 for all groups, p > 0.05 for all values (t-test).

**Table 4. Body weight in backcrossed (generation five) animals of different ages, study A part 2**

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Body weight (g) at Age (months)</th>
<th>3m</th>
<th>6m</th>
<th>7m</th>
<th>8m</th>
<th>9m</th>
<th>10m</th>
<th>11m</th>
<th>12m</th>
<th>13m</th>
<th>14m</th>
<th>15m</th>
</tr>
</thead>
<tbody>
<tr>
<td>Testatin -/-</td>
<td></td>
<td>26.5 ± 0.8</td>
<td>29.9 ± 1.1</td>
<td>34.6 ± 1.2</td>
<td>37.8 ± 1.2</td>
<td>38.3 ± 1.5</td>
<td>41.2 ± 1.1</td>
<td>42.6 ± 1.3</td>
<td>43.5 ± 1.1</td>
<td>43.6 ± 1.2</td>
<td>42.9 ± 1.5</td>
<td>41.8 ± 1.7</td>
</tr>
<tr>
<td>Testatin +/-</td>
<td></td>
<td>28.7 ± 0.5</td>
<td>33.0 ± 0.6</td>
<td>38.2 ± 0.8</td>
<td>41.4 ± 1.0</td>
<td>42.1 ± 1.3</td>
<td>42.1 ± 1.1</td>
<td>45.0 ± 1.1</td>
<td>45.3 ± 1.2</td>
<td>45.8 ± 1.3</td>
<td>46.2 ± 1.6</td>
<td>46.5 ± 2.0</td>
</tr>
</tbody>
</table>

All values are means ± standard errors of means, n = 10, p < 0.05 for values with a star, all other values p > 0.05 (t-test).

**Table 5. Percent body fat (DEXA) in backcrossed (generation five) animals of different ages, study A part 2**

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Percent Body Fat at Age (months)</th>
<th>6m</th>
<th>9m</th>
<th>12m</th>
<th>15m</th>
</tr>
</thead>
<tbody>
<tr>
<td>Testatin -/-</td>
<td></td>
<td>23.0 ± 2.7</td>
<td>33.3 ± 2.7</td>
<td>35.0 ± 1.6</td>
<td>35.3 ± 2.6</td>
</tr>
<tr>
<td>Testatin +/-</td>
<td></td>
<td>25.7 ± 1.7</td>
<td>34.4 ± 2.7</td>
<td>33.5 ± 1.8</td>
<td>39.2 ± 2.0</td>
</tr>
</tbody>
</table>

All values are means ± standard errors of means, n = 10, p > 0.05 for all values (t-test).
Table 6. Body weight in fully backcrossed male mice, study A part 3

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Body weight (g) at Age (months)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1m</td>
</tr>
<tr>
<td><strong>Testatin -/-</strong></td>
<td>14.39 ± 0.402 (n = 10)</td>
</tr>
<tr>
<td><strong>Testatin +/-</strong></td>
<td>14.7 ± 0.601 (n = 12)</td>
</tr>
</tbody>
</table>

All values are means ± standard errors of means, n for each group is indicated after each mean value, p > 0.05 for all values (t-test).

The reproductive tissue specific cystatin subgroup of genes: expression during gonadal development in wild type and testatin knockout animals (Paper II)

A conclusion from publication I was that redundancy between the Testatin/Cres subgroup genes may explain the lack of abnormal phenotype in the testatin knockout mice. Previously, only testatin expression had been reported in fetal testis. The aim of this study was to evaluate if any other of the subgroup genes is expressed during gonad formation as well as to analyse the subgroup expression in testatin knockout mice. mRNA expression of all eight reported Testatin/Cres subgroup genes was evaluated with real-time PCR in fetal testis and ovary, in adult testis and with in situ hybridization in fetal testis.

We found that three of the subgroup genes, CstTE-1, CstSC and Cres, were expressed in wild type fetal testis together with testatin (Fig. 2. Paper II), in the same cellular compartment (Sertoli and germ cells) (Fig. 4. Paper II). None of the genes was upregulated in the fetal testis as a response to the lack of testatin (Fig. 3. Paper II). However, we found that one of the Testatin/Cres subgroup genes, CstE2, was strongly downregulated in adult testis, but not in adult epididymis, in testatin knockout mice (Fig. 5. Paper II). This indicates a possible regulatory connection between testatin and CstE2 in adult testis.

Evolution and human tissue expression of the TCS genes, a reproductive tissue specific subgroup of the family 2 cystatins (Paper III)

We have in this study performed evolutionary studies of the Testatin/Cres subgroup based on gene searches in genomes from eleven species. We have localized the evolutionary origin of the Testatin/Cres subgroup to the split of marsupial and placental mammals (Fig. 6). A model for the evolution of these genes illustrates that they constitute a dynamic group of genes, which has undergone several gene expansions (Fig. 4. Paper III) and we find indications of a high degree of positive selection, in striking contrast to what is seen for the classical cystatin C (Fig. 2. Paper III).
We suggest a new nomenclature for the Testatin/Cres subgroup (TCS) genes, based on their syntenic gene location on mouse chromosome 2 (Table 3. Paper III) and show with phylogenetic relations that the TCS genes are clustered into three original groups, a testatin (TCS7), a Cres (TCS5) and a CstL1 (TCS1) group (Fig. 3. Paper III). We have further characterized the expression patterns of all human members of the subfamily. Of a total of nine identified human genes, four express putative functional transcripts with a predominant expression in the male reproductive system (Fig. 5A and B Paper III), which is in line with a suggested role in reproduction. The remaining five genes do not seem to encode functional transcripts.

Figure 6. Evolutionary origin of the TCS genes. The original picture “summary of paleontological constraints on vertebrate phylogeny” was created by Philip C J Donoghue and Michael J Benton, Department of Earth Sciences, Bristol UK (http://www.fossilrecord.net). Dates are minimum constraints in Million years. The evolutionary origin of the TCS genes is indicated in the figure, as well as the mammalian clades, Eutheria, Marsupial and Monotreme and the Eutherian division into Afrotheria, Xenerthera and Boreoeutheria.
The Testatin/Cres subgroup genes in disorders of sexual development (unpublished project B)

Finding new candidate genes in testis development is important in order to improve diagnostics for patients with disorders of sexual development (DSD). Because of specific expression in mouse fetal testis of four Testatin/Cres subgroup genes we consider that the human TCS genes in general, may be potential candidate genes for involvement in testis formation.

Our group has during 10 years collected a rare group of XY female patients with gonadal DSD (complete 46,XY gonad dysgenesis). Fifty percent of these patients were previously diagnosed with mutations, deletions or duplications in genes known to be involved in testis development, or with cryptic X chromosome mosaicism (Barbaro 2008). However half of the patients still remain without a molecular diagnosis. In this study, we have screened these patients for mutations in the four Testatin/Cres subgroup genes that expressed functional transcripts in human testis (Fig. 5A. Paper III).

Exons and promotor regions in TCS2/CST11, TCS5/CRES, TCS7A/CST9 and TCS7B/CST9L were sequenced in eight XY female patients with gonad dysgenesis. We detected normal SNP:s as well as three variants or mutations previously not reported as common SNP:s (Ensembl 2008). One new variant (−652C>T) was found in the promotor region of TCS7B/CST9L. Another new variant (+13A>C) was found in the 3’ untranslated region of TCS7A/CST9 and the new third variant (c.5C>T) was found in the coding region of TCS7A/CST9, causing a S2L amino acid change in the predicted signal peptide. All new variants were found in different patients. Further investigations are needed to address the functional significance of these findings.

Production of recombinant proteins (unpublished project C)

Our aim was to produce recombinant testatin proteins that can be used in future projects such as crystallographic studies, production of antibodies as well as in functional studies.

Mouse testatin and cystatin C protein (as positive control) were produced in eukaryotic cells (EBNA 293). Secreted protein was analyzed with Western blot (Fig. 7). We detected a prominent band at the expected molecular weight of testatin and an additional band at higher mw. Cystatin expression was stronger with additional bands at higher and one additional band at lower mw probably due to post translational modifications. Future protein quantification and protein secretion efficiency studies will have to be performed to evaluate the amount of protein produced in this system.
Prokaryotic testatin protein is currently produced in collaboration with professor Ylva Lindqvist and collaborators at MBB, KI for further crystallographic studies. Mouse and human testatin were first expressed in the cytoplasm of *E.coli*, which resulted in very low amounts of soluble protein. Cystatin C was previously reported to be successfully produced in the periplasma of *E.coli* (Abrahamson et al. 1988), where, for example, disulphide bridges are allowed to form properly. We have recloned the human and mouse testatin sequences into a vector system that will enable periplasmatic expression in *E.coli*. Production is currently under evaluation.

Production of antibodies to the human Testatin/Cres subgroup genes is ongoing as a part of the HPR (Human Protein Resources) project and we have so far received a CRES antibody.

**Figure 7.** Western blot showing secreted Testatin and Cystatin C from transfected EBNA 293 cells. Supernatant was collected at three different time points, day two after transfection (D2), day four after transfection (D4) and day six after transfection (D6), from two different testatin transfections (Testatin 1 and Testatin 2) and one Cystatin C transfection. His tagged protein was purified and run on a western blot for detection with anti-his antibody. Testatin shows one major form at the expected size 15 kD and one bigger form slightly under 25 kD. Cystatin C shows two major forms, one of 15 kD as expected and one slightly below 15 kD, plus additional forms with higher molecular weight. Protein secretion increases with time. Molecular standard (M) in kD. Negative control (N) is supernatant from untransfected EBNA 293 cells.
DISCUSSION

The Testatin/Cres subgroup genes will from here on be referred to as the TCS genes, as proposed in publication III (Table 7).

<table>
<thead>
<tr>
<th>TCS name Mouse name1</th>
<th>Mouse name 2</th>
<th>Mouse name 3</th>
<th>TCS name Human name 1</th>
<th>Human name 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCS1 CstL1</td>
<td>Cres4</td>
<td></td>
<td>TCS1 CSTL1</td>
<td></td>
</tr>
<tr>
<td>TCS2 CstE1</td>
<td>Cres2</td>
<td>Cst11</td>
<td>TCS2 CST11</td>
<td></td>
</tr>
<tr>
<td>TCS4 CstTE1</td>
<td>Cres3</td>
<td>Cst12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TCS5 CstSC</td>
<td>Cres</td>
<td>Cst8</td>
<td>TCS5 CST8</td>
<td>CRES</td>
</tr>
<tr>
<td>TCS6 CstT</td>
<td>Cst13</td>
<td>Cst9</td>
<td>TCS6 CSTT</td>
<td></td>
</tr>
<tr>
<td>TCS7 Testatin</td>
<td></td>
<td></td>
<td>TCS7A CST9L</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TCS7B CLM</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TCS7C</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TCS7D</td>
<td></td>
</tr>
<tr>
<td>TCS8 CstE2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Putative functions of the TCS genes in testis development

The high expression of TCS7 immediately after Sry in mouse fetal testis, strongly suggests a role in testis formation. However we could in publication I show that TCS7 alone is not necessary for normal mouse testis formation. This was additionally confirmed by the report of a study of another testatin knockout mouse that also showed a normal phenotype (Hasegawa et al. 2006). We could further show in publication II that this phenotype may be explained by functional redundancy with three other members of the TCS genes namely TCS3, TCS4 and TCS5. mRNA for all three genes was detected in fetal testis in the same cellular compartment as TCS7 (Fig. 2 and 4 Paper II).

Since mRNA translation is a highly regulated process, we searched for miRNA recognitions sites (Zamore and Haley 2005) in the TCS genes to investigate the possibility of regulation through degradation. All currently known miRNA sites (approximately 500) were screened in (www.targetscan.org) in all mouse and human TCS genes (Table 3 Paper III). Only one miRNA recognition site was identified in one of the genes namely human TCS1 (unpublished data). Interestingly, the human TCS1 gene turned out to produce a non functional mRNA transcript in adult human testis (Fig. 5B. Paper III). One should not exclude the possibility of finding new miRNA sites that may be present in the TCS genes. However, at this stage a likely conclusion is that the TCS genes probably are not regulated by miRNAs. Protein localization as well as target specificity studies will be important to further investigate the possibility of redundancy between TCS3, TCS4, TCS5 and TCS7 in fetal mouse testis.
We know from publication III that the TCS genes are present only in placental mammals (eutherians). The minimum constraint for the eutherian / marsupial split is estimated to 124.3 MYA (Benton et al. 2008). We have found the TCS genes in all eutherian clades investigated, such as Afrotheria (elephant), Xenarthra (armadillo) and Boreoeutheria (cow, horse, dog, mouse, rat, rhesus and human). This suggests that the evolutionary origin of the TCS genes was before the split of the Boreoeutheria / Xenarthra-Afrotheria clades (Fig. 6. Publ III). The minimum constraint for the Boreoeutheria / Xenarthra-Afrotheria split is estimated to be 62.5 MYA (Benton et al. 2008). Theoretically, the origin of the TCS genes could be earlier than the marsupial placental mammal split (124.3 MYA). In this case a selective pressure in the placental mammals different from the marsupials lead to an expansion of the TCS genes only in the placental mammals while the TCS genes were lost in marsupials. This scenario leads to interesting speculations related to fundamental differences in early sexual development between marsupial and placental mammals.

A search in the literature was performed to identify differences in the pathways leading to ovary or testis formation particularly between marsupials and placental mammals. It has been known for a long time that sex determination signals across vertebrates show a great diversity, however biologists have traditionally believed that the downstream signals might be more conserved, since the structural organization of the adult testis and ovary is very similar across all vertebrates. However, emerging data indicate that downstream pathways and morphogenetic events leading to the development of a testis or an ovary are surprisingly diverse and not as conserved as one earlier believed (reviewed in Yao and Capel 2005).

The role of estrogen is particularly interesting. Estrogen seems to play an important and conserved role in ovary development in most egg-laying vertebrates as birds, fish and alligators. Strikingly, in placental mammals, where embryos are exposed to estrogens from the maternal circulation, the initial prenatal steps of sex determination are not sensitive to estrogen (Britt et al. 2000; Couse et al. 1999). Marsupials show a middle way solution were ovary and testis development is initiated first after birth (Mackay et al. 2004) at the same time as sex determination is clearly estrogen dependent. Administration of estrogen to newborn tammar wallaby (Macropus eugenii) causes complete male to female sex reversal (Coveney et al. 2001). In this perspective it is tempting to imagine that the evolution of a prolonged pregnancy requiring a placenta together with a high estrogen environment for the embryo sparked the development of an alternative estrogen independent female and male pathway in placental mammals. It also provides the basics to speculate about a unique role for the TCS genes only in placental mammals. So far expression of the TCS genes has only been investigated in fetal
mouse gonads (Frygelius et al. 2007). Detection of expression of TCS genes in fetal testis from additional eutherian mammals could further support this hypothesis.

In unpublished project B we have started to analyze the role of the TCS genes in human testis development. The four human TCS genes that expressed functional transcripts in adult testis were sequenced in XY individuals with gonadal DSD. We know from publication II that there may be redundancy between the mouse TCS genes in fetal testis. In the light of this, it would be important to also search for larger deletions comprising several TCS genes in these patients. At the same time phenotypic differences between mouse models and human individuals with deletions /mutations in the same genes are sometimes reported (Val and Swain 2005) and one cannot rule out the possibility that only one gene plays a dominant role in humans, especially since we know from the evolutionary study that the TCS genes are rather different between mice and humans. Interestingly, we have found two different mutations in the TCS7A gene in two different patients, one leading to an amino acid substitution in the signal peptide and one in the non coding 3’ end of the transcript. Further analyses are necessary to understand the biological relevance of these mutations.

Other putative functions of the TCS genes

Evolutionary biologists have typically invoked two types of selective forces that shape the evolution of species. One is purifying selection, which favors the conservation of existing phenotypes. The second is positive selection, which promotes the emergence of new phenotypes (Vallender and Lahn 2004). The evolution of the TCS genes (Fig. 4. Paper III) reveals several gene expansions, which suggests a positive selection. One of the greatest gene expansions seems to have taken place before the split of the Boreoeutheria / Xenarthra-Afrotheria clades with five new TCS genes (TCS1, 2, 5, 6, 7), divided into three original groups, the TCS5 group, the TCS7 group and the TCS1 group (Fig. 4. Paper III). The second expansion took place after the split of the Boreoeutheria-Xenarthra/Afrotheria clades, giving TCS3 and TCS4, in the TCS7 group. More recently, rodents and primates show the highest degree of gene expansions, TCS8 in rodents and a TCS7 expansion in primates (TCS7A-D), suggesting species specific positive selection. Most of the TCS genes show a high conservation in the C-terminal PW consensus site consisting of the amino acids proline and tryptophan. However, three of the TCS genes namely TCS4, TCS6 and TCS7 show a high degree of variation in the PW-site between different species. This is especially true for TCS4, although this diversity is limited to non-polar amino acids such as proline, valine and alanine (Fig. 2. Paper III) there has not been a selection for preserving the tryptophan. The conserved tryptophan in the PW consensus site has been shown to be important for function (Bjork et al. 1996). Therefore the strong conservation in the PW consensus site for most of the TCS genes suggests that this consensus site may be important for their function. TCS4, TCS7A and TCS8
on the other hand may have gained a different specificity since the tryptophan is not conserved. All together, there has been a dynamic evolution of the TCS genes suggesting a strong positive selection.

Comparative genomic analyses reveal that genes involved in reproduction and host pathogen interactions generally are under strong positive selection (Gibbs et al. 2004; Koonin and Rogozin 2003; Vallender and Lahn 2004). We report a strong male reproductive tissue specific expression of the human TCS genes, consistent with what is previously reported for mouse TCS genes. However, two of the human TCS genes, TCS2 and TCS7A show a broader tissue expression. Both are expressed in the immune system related organ thymus (Fig 5A Paper III). Interestingly, the recombinant human TCS2 protein from both the shorter and longer transcripts, has been shown to be antibacterial (Hamil et al. 2002). Some proteins such as β-defensins have been reported to evolve under positive selection for involvement in both male reproduction and host pathogen interactions, both as fertility promoters and anti microbial proteins (Radhakrishnan et al. 2007). This could as well be the case for the TCS genes.

Individual TCS genes may have evolved other functions as well, such as TCS7A, which shows a strong expression in kidney. Moreover, mouse TCS5 was previously reported to be expressed in anterior pituitary gonadotroph cells specifically in the secretory granules. Together with its inhibition of prohormone convertase 2 (Cornwall et al. 2003) this suggested that TCS5 plays a role in the regulation prohormone convertases and the hormones that they regulate (Sutton et al. 1999; Sutton-Walsh et al. 2006). We could not find any expression of TCS5 in human pituitary. This could be due to that only small amounts of the total pituitary that we used contained TCS5 mRNA or that expression of TCS5 in the pituitary is species specific.

We have reported that only four out of totally nine identified human TCS genes are actually expressed in the 23 different human tissues we have investigated (Fig. 4, 5A and 6B Paper III). Mouse has previously been reported to have eight identified TCS genes, all with putative functional transcripts expressed mainly in the male reproductive tract (Frygelius et al. 2007) (Fig. 4. Paper III). Mice show a highly polyandrous mating behavior, while humans show a monandry mating behavior. Postcopulatory sperm competition is a key aspect of sexual selection and it is well established that mating behavior determines the intensity of sperm competition in mammalian species (Dorus et al. 2004; Ramm et al. 2008). A polyandrous mating behavior favors larger testis size, larger seminal vesicles and higher sperm count, compared to monandry mating behavior. Future studies, comparing the number of functional TCS genes in promiscuous versus less promiscuous rodent or primate species could highlight if there is a correlation between number of functional TCS genes and degree of sperm competition. This would further support a function of these genes in male reproduction.
We showed in publication I that *testatin* deficient mice have normal spermatogenesis and fertility. Interestingly, adult *testatin* deficient mice also had a strong downregulation of TCS8 in adult testis. This points at an interregulatory system within the TCS genes but it also shows that TCS8 and TCS7 are not required in the testis for normal fertility. However TCS8 was not downregulated in epididymis, which might be sufficient for normal fertility. We also know through the evolutionary study of the TCS genes that TCS8 only exist in rodents. This further emphasizes the interspecies differences that we have observed. It might as well indicate that the interregulatory phenomena we observed are specific for the rodent TCS genes.

The expression profiles of the mouse TCS genes are highly overlapping in adult testis. Several genes are expressed in Sertoli cells and in the caput part of epididymis (Table 3 Paper II). In contrast, TCS5 and TCS4 proteins have only been detected in different types of spermatides and different compartments in spermatazoa which have led to speculations that they possess unique roles during spermatogenesis and sperm maturation (Li et al. 2005; Yuan et al. 2007). The lack of conservation in the PW site for TCS4 and its unique diversity in this site between different species (Fig. 2 Paper III) might as well hint at a specialized role. In conclusion, the mouse TCS genes might have evolved to play both unique and redundant functions in spermatogenesis and sperm maturation. Generation of conditional knockout mice for all TCS genes would prove further if they play a role in male fertility.

**Spatial localization of the TCS proteins**

The TCS genes are putative extracellular proteins due to the existence of a signal peptide, although emerging data indicate that intracellular functions may exist as well. Alternative start codons after the signal peptide exist for both TCS5 (Cornwall et al. 1999) and TCS7 (Tohonen et al. 1998), and mouse TCS4 protein has been detected both intra and extracellularly. TCS4 protein was detected in the cytoplasm of Sertoli cells and epithelial cells of the caput epididymis. TCS4 protein was also secreted in the lumen of caput epididymis which made the authors speculate that TCS4 may be secreted by Sertoli cells and epithelial cells in the caput epididymis and taken up by maturing spermatozoa (Li et al. 2005). We have expressed mouse TCS7 protein together with its native signal peptide in EBNA293 cells and we detected rather low amounts of secreted protein compared to Cystatin C. Interestingly, TCS4 protein was detected in a stage dependent manner in Sertoli cells at the same time as its mRNA expression was constant. This suggests post transcriptional regulation of the TCS4 gene (Li et al. 2005). Since TCS7 is also expressed in adult Sertoli cells, post transcriptional regulation may regulate TCS7 protein production. Studying regulation of the TCS genes in eukaryotic cell systems will allow further *in vitro* evaluation of their regulation and processing.
CONCLUSIONS AND FUTURE DIRECTIONS

Future in vivo and in vitro studies will help us to understand the putative functions of the TCS genes in reproduction and/or perhaps in other physiological processes as well.

Data presented in this thesis suggest a role for some of the TCS genes in early sexual development in placental mammals. The in vivo role of the TCS genes was studied in both humans and in mice. An expanded mutational screening in patients with DSD as well as generation of knockout mice for all TCS genes expressed in mouse fetal testis would further clarify if the TCS genes play a role in sexual development in placental mammals.

The evolutionary study performed in this thesis as well as the expression patterns of the TCS genes in general are compatible with a role in male reproduction. To further investigate the in vivo role of the human TCS genes in male fertility, we are currently performing an association study using single nucleotide polymorphisms (SNPs) in the 20p region where all TCS genes are clustered. The distribution of these SNPs will be compared among clinically well characterized males with oligo-/azoospermia and controls. Additionally, a comparison between the number of functional TCS genes in promiscuous versus less promiscuous primate species could be performed to show if there is a correlation between number of functional TCS genes and degree of sperm competition.

Finally, production of recombinant TCS proteins was initiated in this thesis to prepare for in vitro functional studies of the TCS genes. It will be particularly important to generate specific antibodies to all TCS genes expressed in mouse fetal testis and to perform protein localization studies to further evaluate their role in sexual development.

The evolution of the amino acid sequences in all three consensus sites was compared in paper III (Fig. 2. Paper III). Most TCS genes show conservation in the PW consensus site, while a few genes show no or little conservation. Cystatin C on the other hand shows high conservation in all three consensus sites between species which implies that the protease specificity in this active site is under purifying selection. Protease inhibition has been reported for one TCS gene previously, TCS5, which was shown to inhibit the serine protease prohormone convertase 2 in vitro, and not cathepsin B and papain as the classic family 2 cystatins (Cornwall et al. 2003). Additionally, human cystatin C has been reported to inhibit legumain with a different active site than the three consensus site traditionally involved in papain inhibition (Alvarez-Fernandez et al. 1999; Chen et al. 1997). Therefore, solving the 3D structure of the TCS proteins as well as evaluating substrate specificity for each gene will be important to further clarify the relationship between structural features and the functions for the different TCS genes.
ACKNOWLEDGEMENTS

I would like to express my thanks to the following persons:

Anna Wedell, my first main supervisor, for your support, your always excellent scientific judgments, for having trusted me to work often very independently and also for all things I have learned from you, especially in clinical science.

Virpi Töhönen, my second main supervisor, for your encouragement, always strong positive drive and for all valuable and nice discussions.

Katarina Nordqvist, my supervisor, for your support, creativity and for adding your scientific experience in the field.

The group, Tiina Robins for compassion and sharing of many interests, Svetlana Lajic for many lively and nice discussions, Tatja Hirvikoski my expert help in statistics but also for many interesting and nice discussions, Michela Barbaro for your scientific interest and lively personality, Annci Thelander for your wonderful compassion and excellent work, Michael Oscarson for bringing a male aspect to the group and Anna Nordenström for always adding an interesting point of view.

Our collaborators, Lauri Pelliniemi for a very nice collaboration on electron microscopy, Ulrik Kvist with whom I performed the spermatozoa evaluation in the testatin knockout mice and who thought me more about the epididymis and Majid Mohammadi that helped with technical aspects, Timo Pikkarainen for generously have given me EBNA 293 cells and together with Max Schneller gave me professional advice in how to express proteins in an eucaryotic cells system. Ylva Lindqvist, Maria Svärd and Jodie Guy for a very interesting and nice collaboration with prokaryotic expression of Testatin. Lars Arvestad, for an excellent collaboration in the evolutionary studies, a great collaboration in every respect.

Phil Donogue and Mike Benton, for generously letting me to use their figure “Summary of paleontological constraints”, in my thesis.

Moderna museet and John Peter Nilsson for help with the picture “Iru” of Olle Bärtling on the front page.

To all co-workers at the CMM floor 02 for creating a nice and lively enviroment and especially, Beni Amatya for sharing many nice issues of life, Andor, Enikö, Agne, Anna B, Sanna and Annika Säaf for many nice discussions about science and life. Agneta Nordenskjöld and her group for sharing lab space and scientific interests, Magnus Nordenskjöld, Ann Nordgren and Giedre for being especially nice teaching companions at the clinical genetics course, teaching me a lot more about clinical genetics. Barbro Werelius, Anna-Lena Kastman, Christina Nyström and Sigrid Sahlen for technical support. The Hamsten group that entered our floor and brought new energy and joy, Petra, Karin, Anna A and Stina.

The Wine seminar with Jan Zedenius and all members for nice scientific discussions and evenings.

Administrative personal at CMM especially Lennart Helleday for flexible and immediate help with my computer.

The fourth floor in CMB, Christer Höög, and his lab where I spent the first part of my thesis, for providing an exciting research atmosphere and for generosity in all matters. I especially want to thank Daniel Lightfoot, Anna Kouznetsova, Katarina Gell, Hong Wang and Lena Ström for valuable particle help as well as nice discussions.
Camilla Sjögren for your encouragement and many interesting discussions about science and life, as my mentor.

Bioinvent International AB, Lund, especially Lena Danielsson for support and teaching me a lot in science and former Head of Research Roland Carlsson, who is no longer with us, for his support, passion for science and generosity.

Eva Engvall, my supervisor in the Burnham Institute, La Jolla, in the very beginning of my research carrier. For a wonderful time in your lab where a learned a lot about cutting edge science in the international world.

My dear, dear friends, for being there and brighten up my life, Linda, Johanna, Maria, Anna, Sara, Eva, Emmelie, Maria T, Jenny F, Jenny L and Zufan.

Sandra my dear wonderful sister, for always understanding me, being there when I need you and for your wise opinions about everything from science to people.

Mamma och Pappa, for your wonderful support in all matters. Especially a great, great thanks for always helping with baby-sitting when needed, you have saved many experiments and deadlines along the way.

My own wonderful family, Anja and Leila for always cheering me up with your wonderful energy and personalities. Max, special thanks for much valuable scientific advice but most of all for bringing passion, love and meaning to my life.
REFERENCES


Viger, R. S., C. Mertineit, J. M. Trasler, and M. Nemer. 1998. Transcription factor GATA-4 is expressed in a sexually dimorphic pattern during mouse gonadal development and is a


