Title
The role of the DAZ gene family in human germ cell development

Author
Moore, Frederick Lee

Publication Date
2002

Peer reviewed|Thesis/dissertation

The Role of the *DAZ* Gene Family in Human Germ Cell Development

by

Frederick Lee Moore

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Genetics

in the

GRADUATE DIVISIONS

of the

UNIVERSITY OF CALIFORNIA SAN FRANCISCO

and

UNIVERSITY OF CALIFORNIA BERKELEY

Date University Librarian

Degree Conferred: .....

Dedication and Acknowledgements

I first would like to thank God for giving me the strength, determination, and will power to stay in school over the last decade. I truly would be lost without you. I would like to thank my parents for their love, support, and friendship. Everything that I am is because of you. I would like to also thank, Renee Reijo-Pera, for your guidance, friendship, and patience. I know that I am not the easiest person to get along with. I enjoyed the journey that we took together and I cherish all of the memories. I would like to thank my brother, Karl Moore, for his love and support. You are one of my best friends in the world and I proud to call you my big brother. To all of my friends that have supported me over the years, thank you for being there for me. I have nothing but unconditional love for you. All of our dreams will come true someday, if we believe and push on. I would like to thank the members of my qualifying exam committee and thesis committee. You challenge me and I grew from it. Thank you. Last of all, to all of the young men and women that march onward and upward toward the light- you have a helping hand in me.

The Role of the *DAZ* Gene Family in Human Germ Cell Development

by

Frederick L. Moore

Understanding of development of germ cells in men and women is critical to understanding causes of infertility and discovering ways to prevent and treat reproductive disorders. The *DAZ* genes, a cluster of four nearly identical genes on the Y chromosome, are deleted in 10% of infertile men with spermatogenic defects characterized by few sperm (oligospermia) or no sperm (azoospermia). Expression of *DAZ* is restricted to male germ cells, whereas expression of an autosomal homolog, *DAZ-Like,* is restricted to male and female germ cells. To characterize the function of DAZ, we sought to identify DAZ-interacting partners by using DAZ as bait in a two-hybrid screen. Seven genes that encode proteins that potentially interact with DAZ were identified; five were further characterized. Two known genes (human *PUMILIO2*, *PUM2*, and human *BOULE*, *BOL*) identified in our screen that were previously characterized in *Drosophila* provided insight into the molecular function(s) of DAZ and DAZL. In *Drosophila,* the *Pumilio* gene is required to maintain germline stem cells during oogenesis via translational regulation. We showed that DAZ protein forms a stable complex with the domain of PUM2 that is homologs to the domain of *Drosophila* Pumilio that binds RNA and rescues *pumilio* loss of function mutants. PUM2 is expressed predominantly in human embryonic stem cells and germ cells, and the protein colocalizes with DAZ and DAZL in germ cells. Later, at the onset of meiosis, we show that DAZ forms a stable complex with the human BOL

protein. The *Boule* gene in *Drosophila* has been shown to control the passage of germ cells through meiosis by translational regulation. The interactions of DAZ and PUM2 and between DAZ and BOL, along with characterization of other DAZ- associated cofactors are described further. In addition, a stage specific model of DAZ, DAZL and interacting cofactors in human germ cell development is presented.

Table of Contents

Table of Contents (continued)

List of Tables

List of Figures

List of Figures (continued)

List of Figures (continued)

List of Figures (continued)

Chapter 1

Introduction

Introduction

Human germ cells mature into sperm and eggs via a complex developmental pathway. In males, disruption of the function of critical genes in the germ cell development pathway results in azoospermia (no germ cells), oligospermia (reduction in germ cell number), and/or asthenoospermia (low sperm motility). Based on preliminary studies reported by Tiepolo and Zuffardi (1976), genetic factors required for spermatogenesis were thought to exist on the human Y chromosome. In their initial studies, Tiepolo and Zuffardi demonstrated that six azoospermic individuals had microscopic deletions of the long arm of the Y chromosome, distal Yq[1]. In four cases, the fathers of the azoospermic men were also karyotyped and all carried intact Y chromosomes. On the basis of the observation of these *de novo* deletions in azoospermic men, Tiepolo and Zuffardi and subsequent researchers proposed the existence of an *Azoospermia Factor* (*AZF*) on Yq. At this time, three *AZF* regions (a, b, c) exist on the Y chromosome, located at Yq11[2-4]. I will briefly summarize work on three AZF candidate genes, located within these *AZF* regions; thereafter, I will review research conducted on members of the *DAZ* gene family (*DAZ* is a candidate gene in the *AZFc* region) acquired in humans and model organisms. This review will provide the background, necessary to put in context my pursuit to understand the functional role of the *DAZ* gene family in human germ cell development.

The *RBM* genes

The *RBM* gene family was the first Y-linked *AZF* candidate gene isolated from interval 6 of Yq11 by positional cloning[2] (Fig. 1). *RBM* is a multicopy gene family

with an estimated 30-40 members (some of which are pseudogenes), located on both arms of the Y chromosome[2, 5, 6]. *RBM* encodes a protein that contains a highly conserved RRM (RNA Recognition Motif) domain and four tandem repeats, named the "SRGY box" for its high content of Ser-Arg-Gly-Tyr amino acids. The expression of *RBM* is exclusively in the germ cells of the testis, and the localization of RBM is nuclear, suggesting a role in RNA processing. Deletions of *RBM* member(s) in the *AZFb* region were found in infertile men with either severe oligospermia or azoospermia and may be associated with spermatogenesis arrest at meiosis[2, 4, 6-11]. Strong evidence for the function of *RBM* in human germ cell development has come from deletions of most copies of the mouse *Rbm* genes, which result in high levels of abnormal sperm development[12].

The *DFFRY* gene

The *DFFRY* gene maps to the *AZFa* region in Yq11.2.[13] (Fig. 1). *DFFRY* is the human homologue of the *Drosophila* developmental gene, *fat facets (faf),* and was first linked to spermatogenic defects in a report of deletions in the *AZFa* region of three men with azoospermia[13]. Thereafter, an azoospermic man with a *de novo* 4-bp deletion in a splice-donor site, causing an exon to be skipped and protein truncation, was found in the *DFFRY* gene[14]. In *Drosophila, faf* has been shown to be essential for oogenesis[15] and is a member of a gene family that encodes enzymes that remove ubiquitin from protein-ubiquitin conjugates[16]. In humans, *DFFRY* is expressed in all adult and embryonic tissues, including the testis. Interestingly, the mouse *Dffry* gene is expressed only in the testis and maps to the Sxr^b region of the Y-chromosome[13], where deletions

have been shown to be associated with early spermatogenic failure with loss of germ cells at meiosis[17]. The observation in humans suggest that the *DFFRY* gene may not play an exclusive role in germ cells, yet may have an important impact on spermatogenesis.

The *CDY* genes

The *CDY* gene family contains multiple copies that are between deletion intervals 5L and 6F on Yq11[18] (Fig. 1). These genes encode proteins that contain a chromatin-binding domain and a catalytic domain and have an autosomal homologue, *CDYL* (*CDY-Like*) on chromosome 6[18]. Expression of *CDY* and *CDYL* differ; while *CDY* is testis specific, *CDYL* is expressed ubiquitously[18]. The function of *CDY* remains to be determined, but several copies are located in both the *AZFb* and *AZFc* regions. Thus, based on the location of several copies and testis-specific expression, the *CDY* family of genes may be required for spermatogenesis in men.

The *DAZ* gene family

The *DAZ* genes were identified in a search for genes that mapped to the long arm of the Y chromosome and may be implicated in spermatogenic failure[3]. These genes map to the *AZFc* region (also referred to as the *DAZ* region), which is in interval 6 of Yq11 (Fig. 1). Four copies of the *DAZ* genes are located in *AZFc*. Men with deletions of these genes have defects in spermatogenesis that are characterized by decreased sperm count (either few or no mature sperm). Testicular biopsies indicate that men with deletions of the *DAZ* region have germ cells that are arrested at the pachytene stage of meiosis or completely lack all germ cells, including spermatogonial stem cells, resulting

in the presence of only somatic cells in their testicular tissue[3, 10, 19]. The variation in phenotypes observed with similar deletions of the *DAZ* region suggests that *DAZ* may function early in the germline stem cell population and again at meiosis. The *DAZ* gene, similar to *CDY*, has an ancestral autosomal homologue, *DAZL* (*DAZ-Like*), which is located on chromosome 3p24[20, 21]. Both the *DAZ* and *DAZL* genes encode proteins that contain a domain capable of binding RNA, the RRM (RNA Recognition Motif) domain[22]. Like other proteins with an RRM domain, the human DAZ and mouse Dazl proteins have been shown to bind RNA[23-25]. In addition to the RRM domain, DAZ protein family members also contain 7-24 repeats of a DAZ domain, a domain of 24 amino acids that is rich in N, Y, and Q residues[3, 20]; the DAZL protein contains one DAZ repeat[26]. The expression of *DAZ* and *DAZL* is restricted to germ cells; *DAZ* is expressed in males, and *DAZL* in males and females[27, 28]. Both DAZ and DAZL proteins have been detected in the primordial germ cells (PGCs) of embryonic testis as early as 20 days and have been shown to localize to specific compartments during spermatogenesis in the adult testis[28, 29]. In PGCs and spermatogonia, DAZ and DAZL proteins have been shown to localize to the nucleus and cytoplasm; they then become restricted exclusively to the cytoplasm at the onset of meiosis in spermatocytes[28]. In embryonic ovaries, DAZL protein has been detected in the PGCs of 20 day fetal ovaries and has been shown to localize to the cytoplasm of adult oocytes[27, 29].

Evolution of *DAZ* & homologues in model organisms

The Y-linked *DAZ* gene cluster likely arose from the autosomal *DAZL* gene through transposition, amplification, and pruning[26, 30]. The Y-linked *DAZ* gene cluster

arose recently, based on its presence in old world monkeys and its absence in new world monkeys and other mammals[21, 31-33]. Substantial studies from model organisms suggest that the *DAZ* and *DAZL* genes likely function in germ cell development and are discussed below. In the mouse, a single *DAZ* homologue, *Dazl*, maps to chromosome 17 and is expressed in the germ cells of male and female mice[32, 33]. The loss of function of *Dazl* in *Dazl* null mice results in a reduced number of germ cells and complete absence of gamete production in both male and female mice[34]. Transgenic experiments that introduced human *DAZ* and *DAZL* genes onto a *Dazl* null background demonstrated the ability of the human genes to partially rescue maturation of germ cells in spermatogenesis[35, 36]. The presence of both transgenes resulted in a greater number of early germ cells and enabled prophase spermatocytes to be produced. However, these germ cells failed to promote differentiation into mid to late pachytenes[35, 36].

The role of the *DAZ* and *DAZL* gene family in germ cell development has also been demonstrated in *Drosophila*, *Xenopus*, and *C. elegans* (Table 1). In *Drosophila*, the *DAZ* homologue, *boule*, is expressed exclusively in the testis[37]. Loss of *boule* function in *Drosophila* results in a phenotype that shares similarities with *DAZ* deletions in humans. In homozygous *boule* mutant flies, spermatocytes are formed, but they fail to undergo meiotic divisions[37]. No meiotic figures are observed, and the products of meiosis, 64-cell spermatids cysts, are absent[37]. Likewise, in many men with deletions of the *DAZ* region, the postmeiotic stages of spermatogenesis are rare or absent and spermatogenesis is completely or nearly completely arrested at the end of prophase in spermatocytes[3, 19]. Evidence from *Drosophila* suggests that in flies, Boule protein

functions to regulate entry into meiosis[38]. *Boule* has been shown to interact genetically with *twine*, which is a Cdc25-type phosphatase, and evidence suggest that the Boule protein activates translation of the *cdc25* transcript[38].

Interestingly in *Xenopus*, the *DAZ* homologue, *Xdazl*, which is expressed in the early embryo and in the germ cells of males and females[39], can function in *Drosophila*. When *Xdazl* cDNA was introduced into *boule* mutant flies, the defect in meiotic entry was partially rescued and the meiotic cell division was completed[39]. This suggests that *Xdazl* can function in meiosis. In addition, however, studies in *Xenopus* indicate a role for *Xdazl* in early germ cell development; *Xdazl* RNA was identified as a component of germ plasm, a region of specialized cytoplasm that is rich in RNAs, RNA-binding proteins, and mitochondria, which is required for PGC formation. *Xdazl* has been shown to be critical for the differentiation of PGCs[40]. The above observations strongly suggested that the *Xdazl* gene may function in early germ cell development and again at meiosis; however, as indicated below in Chapter 3, another gene in the *DAZ* gene family has now been identified that is much more closely related to fly *boule* and may be the functional family member that promotes meiosis under normal conditions. The conservation of *DAZ* homologues through evolution also extends to *C. elegans*, urodeles, and zebrafish, and the role of *DAZ* homologues in these organisms indicates a function in early germ cell development and/or at meiosis[41-43].

DAZ-associated cofactors & RNA targets

With the identification of cofactors that interact with DAZ and DAZL proteins the process of elucidating the role of DAZ and its associated proteins at the molecular level has begun. Two novel genes, *DAZAP1* (*DAZ-Associated Protein*) and *DAZAP2*, encode proteins that interact with DAZ and DAZL proteins[44]. *DAZAP1* encodes a novel RNA-binding protein that is expressed most abundantly in the testis, and *DAZAP2* encodes a ubiquitously expressed protein with no recognizable functional motif. Both DAZAP1 and DAZAP2 proteins bind similarly to DAZ and DAZL through the DAZ repeat[44]. Also of interest, DAZ and DAZL proteins can heterodimerize with each other, and homodimerize to themselves[44, 45].

The pursuit of RNA substrates to which the DAZ and DAZL proteins bind has also begun. One candidate RNA substrate for mouse Dazl protein is the *Tpx-1* transcript[24]. *Tpx-1* mRNA expression is reported to begin at the pachytene stage and persist throughout spermatogenesis to the elongating spermatid stage[46-48]. Tpx-1 protein is essential for the progression of spermatogenesis and reportedly functions in attachment of spermatogenic cells to Sertoli cells[46]. We might expect, if the *Tpx-1* transcript is a legitimate target of *Dazl* that misregulation of the *Tpx-1* transcript by the Dazl protein could disrupt cell-cell interaction and cause spermatogenic defects. Another potential RNA substrate for Dazl protein is *Cdc25a* and *Cdc25c* transcripts[24, 25], which are mammalian homologues of a meiosis-specific phosphatase that the *DAZ* homologue in *Drosophila*, *boule*, is believed to regulate[38]. Various other RNA substrate candidates for Dazl also exist, but future experiments that address the

requirement and phenotypes associated with improper regulation is necessary to understand their significance.

The objective of my thesis is to understand the role of the *DAZ* gene family in human germ cell development. My approach is to identify and characterized proteins that interact with DAZ protein to further elucidate its molecular function through it associated cofactors.

Table 1. *DAZ* homologs: Expression and null phenotypes

Genes		Expression			Phenotypes[#]
		PGC/ Germ plasm	Testis	Ovary	
Human:	*DAZ*	+	+	-	Early GC defects & meiotic arrest
	DAZL	+	+	+	unknown
	*BOL**	-	+	-	unknown
Mouse:	*Dazl*	+	+	+	Early GC defects
	*Bol**	-	+	-	unknown
Drosophila:	*Boule*	-	+	-	meiotic arrest
Xenopus:	*Xdazl*	+	+	+	Early GC defects
C. elegans:	*daz-1*	-	-	+	meiotic arrest

+: protein or mRNA expression observed. -: no expression. GC: germ cells; PGC: primordial germ cells. #: mutants or for *Xenopus*, antisense inhibition. *: This thesis.

References

1. Tiepolo, L. and O. Zuffardi, *Localization of factors controlling spermatogenesis in the nonfluorescent portion of the human Y chromosome long arm.* Hum. Genet., 1976. **34**: p. 119-124.

2. Ma, K., et al., *A Y chromosome gene family with RNA-binding protein homology: candidates for the azoospermia factor AZF controlling human spermatogenesis.* Cell, 1993. **75**: p. 1287-95.

3. Reijo, R., et al., *Diverse spermatogenic defects in humans caused by Y chromosome deletions encompassing a novel RNA-binding protein gene.* Nat. Genet., 1995. **10**: p. 383-93.

4. Vogt, P.H., et al., *Human Y chromosome azoospermia factors (AZF) mapped to different subregions in Yq11.* Hum. Mol. Genet., 1996. **5**: p. 933-943.

5. Chai, N.N., E.C. Salido, and P.H. Yen, *Multiple functional copies of the RBM gene family, a spermatogenesis candidate of the human Y chromosome.* Genomics, 1997. **45**: p. 355-361.

6. Elliott, D.J., et al., *Expression of RBM in the nuclei of human germ cells is dependent on a critical region of the Y chromosome long arm.* Proc. Natl. Acad. Sci. U S A, 1997. **94**: p. 3848-53.

7. Kobayashi, K., et al., *PCR analysis of the Y chromosome long arm in azoospermic patients: evidence for a second locus required for spermatogenesis.* Hum. Mol. Genet., 1994. **3**: p. 1965-1967.

8. Najmabadi, H., et al., *Substantial prevalence of microdeletions of the Y-chromosome in infertile men with idiopathic azoospermia and oligozoospermia detected using a sequence-tagged site-based mapping strategy.* J. Clin. Endocrinol. Metab., 1996. **81**: p. 1347-1352.

9. Foresta, C., et al., *Y-chromosome deletions in idiopathic severe testiculopathies.* J. Clin. Endocrinol. Metab., 1997. **82**: p. 1075-80.

10. Pryor, J.L., et al., *Microdeletions in the Y chromosome of infertile men.* N. Engl. J. Med., 1997. **336**: p. 534-9.

11. Baba, K., et al., *Polymerase chain reaction analysis of the Y chromosome long arm in azoospermic patients: lack of the Y chromosome recognition motif (YRRM1) gene.* Int. J. Urol., 1998. **5**: p. 507-9.

12. Mahadevaiah, S.K., et al., *Mouse homologues of the human AZF candidate gene RBM are expressed in spermatogonia and spermatids, and map to a Y chromosome deletion interval associated with a high incidence of sperm abnormalities.* Hum. Mol. Genet., 1998. **7**: p. 715-727.

13. Brown, G.M., et al., *Characterization of the coding sequence and fine mapping of the human DFFRY gene and comparative expression analysis and mapping to the Sxrb interval of the mouse Y chromosome of the Dffry gene.* Hum. Mol. Genet., 1998. **7**: p. 97-107.

14. Sun, C., et al., *An azoospermic man with a de novo point mutation in the Y-chromosomal gene USP9Y.* Nat. Genet., 1999. **23**: p. 429-32.

15. Fischer-Vize, J.A., G.M. Rubin, and R. Lehmann, *The fat facets gene is required for Drosophila eye and embryo development.* Development, 1992. **116**: p. 985-1000.

16. Huang, Y., R.T. Baker, and J.A. Fischer-Vize, *Control of cell fate by a deubiquitinating enzyme encoded by the fat facets gene.* Science, 1995. **270**: p. 1828-31.

17. Burgoyne, P.S., E.R. Levy, and A. McLaren, *Spermatogenic failure in male mice lacking H-Y antigen.* Nature, 1986. **320**: p. 170-2.

18. Lahn, B.T. and D.C. Page, *Four evolutionary strata on the human X chromosome.* Science, 1999. **286**: p. 964-7.

19. Reijo, R., et al., *Severe oligospermia resulting from deletions of the Azoospermia Factor gene on the Y chromosome.* Lancet, 1996. **347**: p. 1290-93.

20. Yen, P.H., N.N. Chai, and E.C. Salido, *The human autosomal gene DAZLA: testis specificity and a candidate for male infertility.* Hum. Mol. Genet., 1996. **5**: p. 2013-7.

21. Shan, Z., et al., *A SPGY copy homologous to the mouse gene Dazla and the Drosophila gene boule is autosomal and expressed only in the human male gonad.* Hum. Mol. Genet., 1996. **5**: p. 2005-11.

22. Burd, C.G. and G. Dreyfuss, *Conserved structures and diversity of functions of RNA-binding proteins.* Science, 1994. **265**: p. 615-621.

23. Tsui, S., et al., *Association of the mouse infertility factor DAZL1 with actively translating polyribosomes.* Biol. Reprod., 2000. **62**: p. 1655-60.

24. Jiao, X., P. Trifillis, and M. Kiledjian, *Identification of target messenger RNA substrates for the murine deleted in azoospermia-like RNA-binding protein.* Biol. Reprod., 2002. **66**: p. 475-85.

25. Venables, J.P., M. Ruggiu, and H.J. Cooke, *The RNA-binding specificity of the mouse Dazl protein.* Nucleic. Acids Res., 2001. **29**: p. 2479-83.

26. Saxena, R., et al., *The DAZ gene cluster on the human Y chromosome arose from an autosomal gene that was transposed, repeatedly amplified and pruned.* Nat. Genet., 1996. **14**: p. 292-9.

27. Dorfman, D.M., D.R. Genest, and R.A.R. Pera, *Human DAZL1 encodes a candidate fertility factor in women that localizes to the prenatal and postnatal germ cells.* Hum. Reprod., 1999. **14**: p. 2531-2536.

28. Reijo, R.A., et al., *DAZ family proteins exist throughout male germ cell development and transit from nucleus to cytoplasm at meiosis in humans and mice.* Biol. Reprod., 2000. **63**: p. 1490-1496.

29. Brekhman, V., et al., *The DAZL1 gene is expressed in human male and female embryonic gonads before meiosis.* Mol. Hum. Reprod., 2000. **6**: p. 465-8.

30. Kuroda-Kawaguchi, T., et al., *The AZFc region of the Y chromosome features massive palindromes and uniform recurrent deletions in infertile men.* Nat. Genet., 2001. **29**: p. 279-86.

31. Delbridge, M.L., et al., *A human candidate spermatogenesis gene, RBM1, is conserved and amplified on the marsupial Y chromosome.* Nat. Genet., 1997. **15**: p. 131-6.

32. Cooke, H.J., et al., *A murine homologue of the human DAZ gene is autosomal and expressed only in male and female gonads.* Hum. Mol. Genet., 1996. **5**: p. 513-6.

33. Reijo, R., et al., *Mouse autosomal homolog of DAZ, a candidate male sterility gene in humans, is expressed in male germ cells before and after puberty.* Genomics, 1996. **35**: p. 346-52.

34. Ruggiu, M., et al., *The mouse Dazla gene encodes a cytoplasmic protein essential for gametogenesis.* Nature, 1997. **389**: p. 73-7.

35. Vogel, T., et al., *Partial rescue of the Dazl knockout mouse by the human DAZL gene.* Mol. Hum. Reprod., 2002. **8**: p. 797-804.

36. Slee, R., et al., *A human DAZ transgene confers partial rescue of the mouse Dazl null phenotype.* Proc. Natl. Acad. Sci., 1999. **96**: p. 8040-45.

37. Eberhart, C.G., J.Z. Maines, and S.A. Wasserman, *Meiotic cell cycle requirement for a fly homologue of human Deleted in AZoospermia.* Nature, 1996. **381**: p. 783-5.

38. Maines, J.Z. and S.A. Wasserman, *Post-transcriptional regulation of the meiotic Cdc25 protein Twine by the Dazl orthologue Boule.* Nature Cell Biol., 1999. **1**: p. 171-174.

39. Houston, D.W., et al., *A Xenopus DAZ-like gene encodes an RNA component of germ plasm and is a functional homologue of Drosophila boule.* Development, 1998. **125**: p. 171-80.

40. Houston, D.W. and M.L. King, *A critical role for Xdazl, a germ plasm-localized RNA, in the differentiation of primordial germ cells in Xenopus.* Devel., 2000. **127**: p. 447-56.

41. Johnson AD, B.R., Drum M, Masi T., *Expression of axolotl dazl RNA, a marker of germ plasm: widespread maternal RNA and onset of expression in germ cells approaching the gonad.* Dev. Biol., 2001. **234**: p. 402-15.

42. Maegawa, S., K. Yasuda, and K. Inoue, *Maternal mRNA localization of zebrafish DAZ-like gene.* Mech. Dev., 1999. **81**: p. 223-26.

43. Karashima, T., A. Sugimoto, and M. Yamamoto, *Caenorhabditis elegans homologue of the human azoospermia factor DAZ is required for oogenesis but not for spermatogenesis.* Devel., 2000. **127**: p. 1069-79.

44. Tsui, S., et al., *Identification of two novel proteins that interact with germ-cell specific RNA-binding proteins DAZ and DAZL1.* Genomics, 2000. **65**: p. 266-73.

45. Ruggiu, M. and H. Cooke, *In vivo and in vitro analysis of homodimerization activity of the mouse Dazl protein.* Gene, 2000. **252**: p. 119-126.

46. Maeda, T., et al., *Molecular cloning of the rat Tpx-1 responsible for the interaction between spermatogenic and Sertoli cells.* Biochem. Biophys. Res. Commun., 1998. **248**: p. 140-6.

47. O'Bryan, M.K., et al., *Identification of a rat testis-specific gene encoding a potential rat outer dense fibre protein.* Mol. Reprod. Dev., 1998. **50**: p. 313-22.

48. Foster, J.A. and G.L. Gerton, *Autoantigen 1 of the guinea pig sperm acrosome is the homologue of mouse Tpx-1 and human TPX1 and is a member of the cysteine-rich secretory protein (CRISP) family.* Mol. Reprod. Dev., 1996. **44**: p. 221-9.

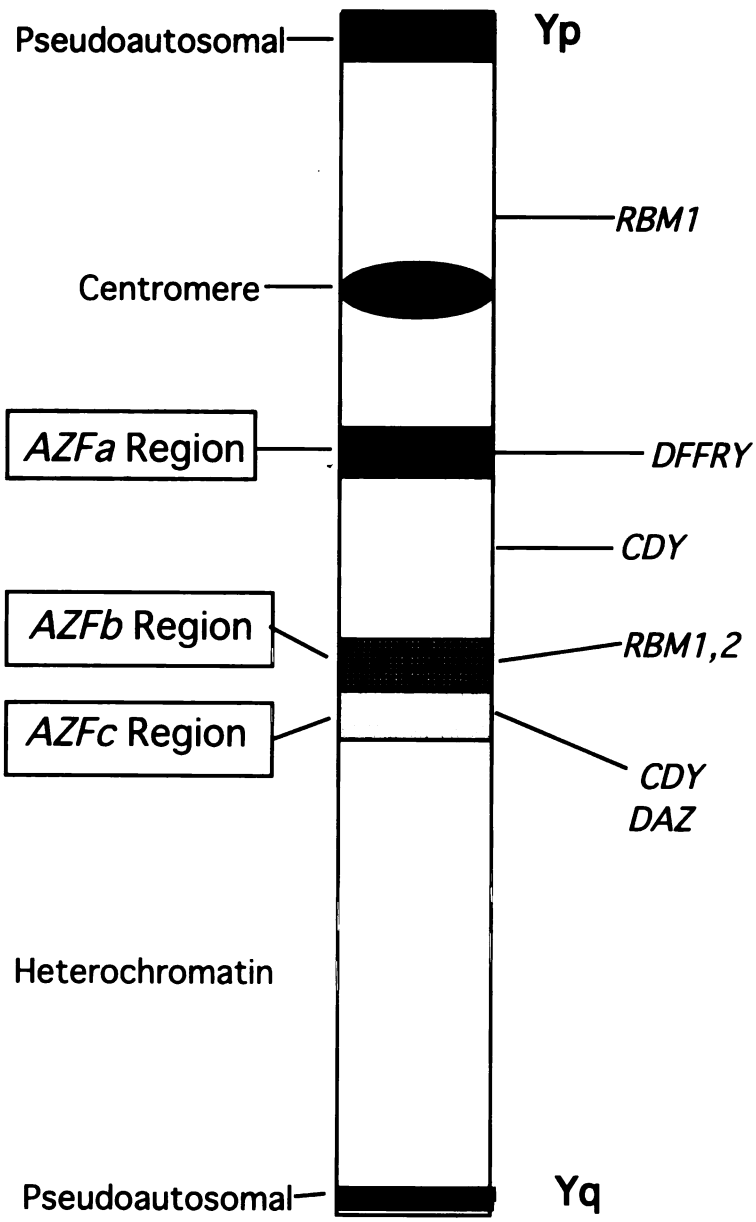

Pseudoautosomal
Yp
RBM1
Centromere
AZFa Region
DFFRY
CDY
AZFb Region
RBM1,2
AZFc Region
CDY
DAZ
Heterochromatin
Pseudoautosomal
Yq

Figure 1

Chapter 2

A Human Homolog of *Drosophila* Pumilio is Expressed in Embryonic Stem Cells and Germ Cells and Interacts with DAZ (Deleted in AZoospermia) and DAZ-Like Proteins

A Human Homolog of *Drosophila* Pumilio is Expressed in Embryonic Stem Cells and Germ Cells and Interacts with DAZ (Deleted in AZoospermia) and DAZ-Like Proteins

Frederick L. Moore[*], Jadwiga Jaruzelska[*†], Mark S. Fox[*], Jun Urano[‡], Meri T. Firpo[*], Paul J. Turek[*], David M. Dorfman[x], and Renee A. Reijo Pera[*]

[*]Department of Obstetrics, Gynecology and Reproductive Sciences; Department of Physiology; Department of Urology; and Programs in Human Genetics and Cancer Genetics; University of California at San Francisco, San Francisco, CA 94143-0546, USA

†Institute of Human Genetics, Polish Academy of Sciences, Poznan, Poland

‡Department of Biochemistry and Biophysics; University of California at San Francisco, San Francisco, CA 94143, USA

[x]Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, 75 Francis Street, Boston, MA 02115, USA

Abstract

Early in development, a part of the embryo is set aside to become the germ cell lineage which will ultimately differentiate to form sperm and eggs and transmit genetic information to the next generation. Men with deletions encompassing the Y-chromosome *DAZ* genes have few or no germ cells but are otherwise healthy, indicating they harbor specific defects in formation or maintenance of germ cells (1-3). A *DAZ* homolog, *DAZL* (*DAZ-Like*), is found in diverse organisms, including humans (4-16) and is required for germ cell development in males and/or females (9, 11, 17, 18). We identified proteins that interact with DAZ proteins in order to better understand their function in human germ cells. Here, we show that PUM2, a human homolog of Pumilio, a protein required to maintain germ line stem cells in *Drosophila* and *Caenorhabditis elegans* (19-24), forms a stable complex with DAZ through the same functional domain of PUM2 required for RNA binding, protein-protein interactions and rescue of *pumilio* mutations in flies. We also show that *PUM2* is expressed predominantly in human embryonic stem cells and germ cells and colocalizes with DAZ and DAZL in germ cells. These data implicate PUM2 as a component of conserved cellular machinery that may be required for germ cell development.

Introduction

All stem cells have potential to differentiate or proliferate mitotically. These potentials must be balanced for the stem cell population to be maintained. If differentiation exceeds proliferation, the stem cell population is not maintained. Evidence in humans suggests that the *DAZ* genes function early in the germ line stem cells. Men with deletions encompassing the Y-chromosome *DAZ* gene cluster have defects in spermatogenesis that are detected initially in the stem cell population (1, 2, 25). These men frequently lack all germ cells, including the spermatogonial stem cells and only somatic cells are present in testicular tissue (1, 2, 25). In addition, expression of the *DAZ* gene and its ancestral, autosomal homolog, *DAZL,* only occurs in germ cells—*DAZ* is expressed in males, *DAZL* in males and females (26-27). Evidence from model organisms also suggests the *DAZ* genes function in germ cell maintenance. The *Xenopus* homolog of *DAZ, Xdazl*, is expressed in a region of the early oocyte that contains the germ plasm that is required for formation and maintenance of the germ cell lineage (7). Inhibition of *Xdazl* leads to loss of the primordial germ cells (17). Finally, the *DAZ* homolog in mice, *Dazl*, is most abundantly expressed in premeiotic germ cells; disruption of this gene causes depletion of germ cells beginning prenatally (18, 26, 28). *DAZ* and *DAZL* homologs may function interchangeably as suggested by the observation that a human *DAZ* transgene can partially rescue a mouse *Dazl* mutation (29). To shed light on how *DAZ* genes might function in human germ cells, we sought to identify proteins that interact with DAZ proteins.

Materials and Methods

Two-hybrid screening of DAZ interacting proteins. The yeast two-hybrid system was used to identify proteins that interact with a DAZ:GAL4 DNA-binding domain fusion protein (Clontech, Inc., Palo Alto, CA). This construct was derived from a cDNA that encodes complete RNA-binding and DAZ repeat domains (pRR102). A library of testis cDNAs fused to the activation domain of GAL4 was screened three times according to manufacturer's instructions. 23 clones were obtained. Those that encoded proteins that interacted with the DNA-binding domain, interacted with unrelated laminin protein, or activated transcription independent of DAZ were not pursued.

Deletion analysis. Deletions of the PUM2 open reading frame were constructed by PCR amplification of overlapping fragments of the *PUM2* cDNA and insertion into the pACT2 vector. Constructs were designed to encode fusion proteins in the correct reading frame, were transformed into yeast with the DAZ:GAL4 DNA-binding domain construct and screened for interaction by X-Gal filter assays. For DAZ deletions, primers were used to clone fragments of *DAZ* cDNA adjacent to the GAL4 DNA-binding domain within the pAS2 vector to produce fusion proteins. Constructs were transformed into yeast with PUM2:GAL4 activation domain construct and screened for interaction as above.

Coimmunoprecipitation of human PUM2 and DAZ. Yeast cells that expressed fusion proteins were pelleted, lysed and resuspended in lysis buffer (50 mM Tris-HCl, pH 7.5, 1 mM EDTA, 150 mM NaCl, 10% glycerol, 1% NP-40) with a cocktail of protease

inhibitors (Clontech, Inc., Palo Alto, CA). The *PUM2* cDNA was cloned in frame with the GAL4 activation domain and a hemagglutinin (HA) tag and used to express protein for immunoprecipitation from yeast. HA antibodies covalently linked to inert beads (BABCO, Inc., Berkeley, CA) were incubated with yeast supernatants, washed with lysis buffer 3 times, resuspended in 1 mM glycine (pH 2.8) and incubated at 30°C to elute proteins. Coimmunoprecipitation from COS-7 cells was accomplished by transfection with a pBud vector (Invitrogen, Inc., Carlsbad, CA) containing the *PUM2* cDNA fused to an HA tag. Cells were cultured for ~48 hours, lysed with lysis buffer as above and incubated with DAZ proteins that were translated in vitro from a pET-29a(+) construct (Novagen, Inc., Madison, WI) that encoded DAZ in frame with an S tag. Cell supernatant with S-DAZ, HA-PUM2 and both proteins were incubated with S-tag beads for 1-2 hours, beads were washed and denaturing at 95 degrees eluted proteins. Coimmunoprecipitation in mouse testis extract was accomplished by transforming BL21-Gold(DE3) pLys bacteria cells (Stratagene, Cedar Creek, TX) with a pET-29a(+) construct (Novagen, Inc., Madison, WI) that encoded DAZ in frame with an S tag. Bacteria cells were sonicated in lysis buffer as above and supernatants containing S-DAZ and containing the S Tag alone but no DAZ-protein fusion were precleared with protein A beads for 30 minutes, and incubated with S-tag beads for 1 hour at room temperature. S-beads were washed, and incubated with testis supernatant for 2 hours at 4 degrees. S-beads were washed, and eluted by denaturing at 95 degrees. Protein electrophoresis and Western were described (26).

Expression analysis. Northern hybridization and RT-PCR were as described (1). Tissues were homogenized in Trizol according to instructions (Gibco-BRL, Inc., Bethesda, MD). One μg of total RNA was reverse-transcribed using oligo-dT primers (Roche, Inc., Indianapolis, IN). 100 ng of cDNA was subjected to RT-PCR. Primers were: PUM1: AAAAACCTGAGAAGTTTGAATTGT and GCAAGACCAAAAGCAGAGTTG; PUM2: ACCAACATTCCTTGGTGAG and ATCAGGACCCCAAGAAGAGG.

Immunohistochemistry. PUM2 antibodies were produced in rabbits with the peptide PNPTANKPLVEEFSNPETQN (Research Genetics, Inc., Huntsville, AL). This peptide is only present in PUM2; it is not in PUM1. PUM2 antisera were collected after 8 weeks. Immunohistochemistry was as described (26). Human testis sections were from a 20-week fetus and a 38-year-old man with a seminoma; sections were not in contact with the seminoma. Human ovary sections were from a 19-week fetus and 37 and 44 year-old women. Mouse sections were from 60-day mice. PUM2 and preimmune antisera were used at dilutions of 1:200 for histology and 1:700 for Westerns. Cell types were assigned by standard criteria (30).

Three Hybrid Assay. Yeast strain L40c, which contains the gene that encodes the LexA-MS2 fusion protein integrated into the chromosome, was transformed with RNA hybrid vector pIII/MS2-2/NRE (31). pIII/MS2-2/NRE contains the binding sites for MS2 coat protein linked to the NRE sequence. To assay for PUM2 interaction with the NRE, *PUM2* cDNA was cloned into the pACT2 vector, which encodes the GAL4 activation domain. In addition, to assay for PUM2/DAZL/NRE complex formation, the *PUM2*

cDNA was cloned into the p412ADH plasmid and transformed with a DAZL fusion construct, containing the GAL4 activation domain in the pACT2 vector. Negative controls are as shown (Fig. 6). X-Gal filter assays were as described (31).

Results

Identification of DAZ-interacting proteins.

We sought to identify proteins that interact with human DAZ proteins via a yeast two-hybrid screen. We identified seven proteins that potentially interact with DAZ (Table 1). These proteins are notable in several ways. (i) Five of the candidate interacting proteins contain domains predicted to bind RNA (PUM2, BOL, DZIP1, DZIP2 and DAZL). DAZ homologs have also been shown to bind RNA and associate with ribosomes (32-34). This suggests that DAZ proteins may function with interacting proteins to regulate translation. (ii) Four of the proteins are encoded by loci shown genetically to be required for fertility in other organisms (PUM2, BOL, hQK3 and DAZL) (1, 18, 20, 35, 36). (iii) Five of the proteins have been shown biochemically to interact with DAZ (BOL, DZIP1, DZIP2, hQK3 (data not shown) and DAZL (32, 37). These observations bolstered our confidence that we had identified legitimate partners.

We further characterized interaction of DAZ with PUM2, a homolog of *Drosophila* Pumilio. Analysis of the PUM2 protein sequence we obtained (GenBank Accession No. AF272350) indicated that the protein was nearly identical to that described in GenBank, Accession No. XM_015812. The exception was that the PUM2 protein encoded by the clone we isolated had an insertion of 79 amino acids at the N-terminus. Overall, human PUM2 protein shares 75% identity with a second human protein, PUM1. PUM2 protein is also 96% identical to a mouse Pum2; 85% identical to the single known member of the family in frogs, XPum; and 35% identical to *Drosophila* Pumilio. The similarity between

PUM2 and other Pumilio proteins was especially striking when we compared the RNA-binding domains of proteins from several organisms including flies (Fig. 1). Remarkably, the human PUM2 RNA-binding domain is 80% identical to that of *Drosophila* Pumilio, the essential functional domain in fly embryogenesis (38).

Interaction of DAZ and PUM2.

To verify interaction of the PUM2 and DAZ proteins that we observed in the two-hybrid system, we pursued biochemical isolation of PUM2 and DAZ complexes. We began with coimmunoprecipitation from yeast strains that expressed DAZ alone, PUM2 alone, and DAZ and PUM2 proteins together (Fig. 2*A*; left panel). We used antibodies that bind a tag on PUM2 to isolate it from yeast supernatants, and assayed for the presence of interacting DAZ protein by Western analysis. As shown in Figure 2*A* (left panel), DAZ protein was coimmunoprecipitated only from supernatant that expressed both DAZ and PUM2. We extended these results to primate COS-7 cells (Fig. 2*A*; middle panel). Antibodies that bind a tag on DAZ were incubated in supernatant from COS-7 cells containing DAZ and PUM2 or each protein alone. Coimmunoprecitation of PUM2 protein was detected by Western analysis only from cell supernatants where both DAZ and PUM2 proteins were expressed (Fig. 2*A*; middle panel). In a similar manner, we also determined whether DAZ and PUM2 proteins interact in the testis (Fig. 2*A*, right panel). To accomplish this, we incubated mouse testis with human DAZ protein coimmunoprecipitated from bacterial extracts expressing DAZ. Association of DAZ and PUM2 occurred only when testis extracts were incubated with DAZ and was not observed in control extracts prepared from bacterial cells that lacked DAZ expression

(Fig. 2*A*, right panel). These data establish that DAZ and PUM2 can form a stable complex.

We narrowed down regions of DAZ and PUM2 that are required for interaction by the yeast two-hybrid system (Fig. 2 *B* and *C*). Analysis of truncated PUM2 proteins indicated that DAZ protein interacts with the RNA-binding region of PUM2, which contains an RNA-binding domain with eight PUF (PUmilio Fbf) repeats (39, 40). Constructs that encoded the entire RNA-binding domain (amino acids 471-876 or 532-876) produced PUM2 peptides that interacted with DAZ protein as well as full-length PUM2 (Fig. 2*B*). In contrast, when residues 795-876, which comprise part of PUF repeat 8 were deleted, interaction with DAZ protein was reduced (Fig. 2*B*). Deletion analysis of DAZ protein was also used to identify regions of DAZ required for PUM2 interaction. Results indicated that PUM2 interacted with the linker region of DAZ between the RNA-binding DAZ repeat domains (amino acids 124-173; Fig. 2*C*).

Expression of *PUM2* mRNA.

For interaction of DAZ and PUM2 to occur in vivo, the genes that encode them must be expressed in the same cellular/subcellular compartments. We examined expression of *PUM2* mRNA and protein and compared it to *DAZ*. *PUM2* mRNA expression was analyzed by both Northern analysis and RT-PCR. As shown by Northern analysis, *PUM2* was highly expressed in adult ovary and testis (Fig. 3*A*). Little expression was detected in other adult tissues (leukocytes, colon, small intestine, prostate, thymus and spleen) (Fig. 3*A*). To extend these results, RT-PCR was used to compare

mRNA expression of *PUM2* with that of *PUM1* in human embryonic stem (ES) cells, various fetal and adult tissues (Fig. 3*B*). RT-PCR analysis indicated that *PUM2* was abundantly expressed in ES cells, fetal and adult ovary and testis; there was little or no expression in other tissues (Fig. 3*B*). In contrast, *PUM1* mRNA was abundant in all tissues (Fig. 3*B*). Given that expression of *PUM2* was greatest in gonads, we asked whether it is expressed in germ cells of the testis, somatic cells or both. To distinguish between these alternatives, expression of *PUM2* and *PUM1* mRNA was compared in tissue samples from men who had germ cells in their testis biopsies and from men who had no germ cells in their testis. Results showed that PUM2 was expressed most abundantly in testis biopsies with germ cells; variation in expression likely reflects variability in germ cell number and type in individual biopsies (Fig. 3*C*). *PUM2* expression was not observed in biopsies from men who had no germ cells (Fig. 3*C*). These results suggested that PUM2 was either expressed in germ cells or required germ cells for expression in somatic cells. In contrast, *PUM1* was expressed in samples from men with and without germ cells (Fig. 3*C*).

Localization of PUM2 protein.

DAZ and DAZL proteins are confined to germ cells (26-27). We generated antibodies specific to PUM2 to compare localization to that of DAZ and DAZL. Western analysis of mouse ES cell, human and mouse testis extracts with PUM2 antisera indicated the presence of a dominant protein band of ~95 kD (lanes 1-3; Fig. 4*A*). Evidence that this represented native PUM2 protein is based upon the observations that this band was not present on Westerns incubated with PUM2 preimmune sera (lanes 4-6) and was

reduced in intensity when PUM2 antisera was preincubated with recombinant PUM2 protein (lanes 7-9). In addition to the dominant ~95 kD PUM2 band, we observed two weaker protein bands of higher molecular weight in human and mouse testis and mouse ES cells that were not present in preimmune sera and were competed by PUM2 peptide. These likely correspond to isoforms of PUM2 generated from the same gene on chromosome 2 that correspond to ESTs whose sequence is deposited in the GenBank database (Accession numbers: XM_015812).

PUM2 antisera, and antisera specific to DAZ and DAZL proteins (26-27), were used to compare expression and cellular localization of these proteins. As previously shown (26-27), both the Y-chromosome encoded DAZ and autosome-encoded DAZL proteins are expressed in germ cells of the human testis—the spermatogonia, early and late spermatocytes and postmeiotic cells (Fig. 4*B*). In females, only DAZL protein is expressed, in the cytoplasm of oocytes (Fig. 4*C*). DAZL expression begins early—prenatally in primordial germ cells of fetal testis (Fig. 4*D*; DAZ is expressed similarly) and fetal ovary (27). The expression of PUM2 mirrors that of DAZ and DAZL: PUM2 protein is most abundant in the cytoplasm and nucleus of spermatogonia but is also present in the cytoplasm of early and late spermatocytes (Fig. 4*E*) and is present in the cytoplasm of oocytes (Fig. 4*F*). Expression of PUM2 was also detected in primordial germ cells of fetal testis (Fig. 4*G*) and fetal ovary (data not shown). As shown, preimmune sera does not cross react with cellular proteins in the germ cells of the testis (Fig. 4*H*), ovary (Fig. 4*I*) and fetal testis (Fig. 4*J*). In addition, use of antisera against

PUM1 results in a different pattern of staining: Germ cells, of all stages, as well as somatic cells demonstrate staining (data not shown).

To further examine expression of Pum2 in a mammal that is amenable to staging of germ cell types and confirm expression in spermatogonial stem cells, we examined the expression of Pum2 in mice (Fig. 5). Expression of Pum2 protein in mice mirrored that of humans in cell type and subcellular localization. The Pum2 protein was abundant in the cytoplasm of adult oocytes (Fig. 5*A*). In the male, Pum2 protein was concentrated in the spermatogonial cells found at the perimeter of spermatogenic tubules of stage I to IX (see inset of Fig. 5*B*; shown is a stage III tubule with A_1 spermatogonia near the perimeter). Immunohistochemistry with preimmune PUM2 antisera did not demonstrate significant signal in oocytes or spermatogonia (Fig. 5 *C* and *D*). Moreover, this pattern of expression was unique and specific as indicated by the use of antisera against Boule, a close relative of DAZ and DAZL (41). Use of these antisera illustrates the differences in expression pattern between premeiotically-expressed DAZ/DAZL/PUM2 proteins and meiotic expression of Boule. The later is not detectable in oocytes (Fig. 5*E*); it is also absent from spermatogonial stem cells; only spermatocytes contained detectable Boule (Fig. 5*F*).

PUM2 and DAZL interact on RNA.

Given our observations that PUM2 and DAZ can form a stable complex and colocalize in the germ cell lineage, we addressed whether PUM2 can interact with DAZL protein and form a complex on RNA. We chose to examine interaction with the NRE

(Nanos Regulatory Element) RNA that is repressed by fly Pumilio. In flies, Pumilio interacts with proteins such as Nanos and Brat on the NRE (42) and represses translation of *hunchback* transcript, to establish the embryonic axis (42-43). To determine whether human PUM2 protein can interact with DAZL and bind the NRE sequence as a complex, we first demonstrated that human PUM2 protein could bind to the NRE (Fig. 6*A*). Then we attached the activation domain of GAL4 to DAZL to assess ability of PUM2 and DAZL to interact and activate transcription of reporter genes, LacZ (Fig. 6*B*) and HIS3 (data not shown). We found that PUM2 and DAZL formed a stable complex, recognized the NRE RNA, and activated transcription (Fig. 6*B*).

Discussion

In model organisms, Pumilio proteins are required for germ line maintenance and differentiation (19-24). Germ line stem cells are not maintained and instead differentiate prematurely in *Drosophila* Pumilio mutants (20). In *C. elegans,* double mutations in the Pumilio homologs, *fbf-1* and *fbf-2,* result in loss of germ line stem cell populations (24). It is proposed that FBF proteins may act to control the maintenance of germ line stem cells, at least partially, by transnational repression of *gld-1,* a gene that may promote meiotic differentiation (24).

Here, we demonstrate that the human DAZ/DAZL proteins can form a stable complex with human PUM2. This observation potentially ties previous data acquired on the function of Pumilio in the germ cells of other organisms to the function of PUM2 and DAZ/DAZL proteins in human germ cells. Evidence suggests human PUM2 may play a

role similar to that of Pumilio in model organisms: First, the RNA-binding region, required to regulate translation during fly embryogenesis and germ line development, is 80% identical to that of humans. Moreover, the RNA-binding region of *Drosophila* Pumilio is the site where cofactors are recruited to PUF repeats 7 and 8 and the sequences C-terminal to facilitate Pumilio function as a translational repressor during embryogenesis (40). Our studies indicate that the RNA-binding domain is required for interaction of PUM2 with DAZ and that deletions that remove part of PUF repeat 8, reduced DAZ/PUM2 interaction. Second, PUM2 protein can bind the NRE sequence, the sequence necessary for translational repression of specific transcripts in *Drosophila;* this suggests that PUM2 may, maintain ability to regulate translation. Finally, the pattern of PUM2 expression in embryonic cells and germ cells suggests that it functions in these cell types.

Thus, based on studies on the *DAZ* gene family and studies in *Drosophila* and *C. elegans* that indicate an essential role of Pumilio in germ cell development (19-24), we suggest that PUM2 may function in maintenance of the human germ cell lineage. PUM2 is expressed throughout the development of the female and male germ cell lineages. Perhaps, then, PUM2 functions as a translational regulator in the germ cell lineage in conjunction with DAZ and DAZL proteins. This hypothesis is consistent with recent reports that indicate translational repression by Pumilio homologs may be an ancient and widespread mechanism for maintenance of germ line stem cells (24, 43). The observation that PUM2 message and protein is expressed in embryonic stem cells is also intriguing, though its role in this cell remains to be explored. With identification of conserved

embryonic and germ cell components, the stage is set to unravel the function of genes such as PUM2 that may be required for maintenance of cells that give rise to mature germ cells in men and women.

Figure legends

Fig. 1. Alignment of RNA-binding domains of Pumilio: Sequences are from human (hPUM1, hPUM2), mouse (mPum1, mPum2), frog (XPum), and flies (DPum).

Fig. 2. Interaction of PUM2 and DAZ proteins. *(A)* The DAZ protein was coimmunoprecipitated with PUM2 when HA beads were incubated with supernatant from yeast that expressed both DAZ and PUM2 fused to an HA tag; shown is DAZ protein detected by Western (left-hand panel). DAZ protein was not coimmunoprecipitated with HA beads from yeast that expressed DAZ or PUM2 alone. Right lane shows yeast supernatant that expressed DAZ as a positive control. Middle panel: PUM2 protein coimmunoprecipitated with DAZ protein from primate COS-7 cells. Western analysis showed that PUM2 protein formed a complex with S-tag:DAZ protein captured from COS-7 cells that expressed PUM2 and DAZ. PUM2 protein was not coimmunoprecipitated from supernatant of COS-7 cells when S-beads were incubated with supernatants from cells with PUM2 or DAZ alone. Cell supernatant that expressed PUM2 was included as a positive control (middle panel: right-most lane). PUM2 protein was coimmunoprecipitated with DAZ from testis extract (right-hand panel). Western analysis showed that PUM2 proteins formed a complex with S-tag:DAZ protein when incubated in mouse testis extract. No PUM2 was detected when S-beads were incubated in mouse testis extract that did not contain the S-tag:DAZ fusion protein. Mouse testis extract included as a positive control (right-hand panel: right most lane). *(B)* Deletions of the PUM2 protein defined minimal domain of PUM2 required for interaction with DAZ.

Eight deletions of PUM2:GAL4 protein were assayed for interaction with intact DAZ in the yeast two-hybrid system. Positions of amino acids on PUM2 constructs are as in GenBank Accession No. AF272350. *(C)* Eleven deletions of DAZ defined minimal region required for PUM2 interaction via the yeast two-hybrid assay. Several constructs truncated the protein near the RRM domain (RNP-2 and -1) and interacted only weakly with PUM2, even though they contain the minimal interaction region. This weak interaction may have been due to improper folding of DAZ due to disruption of RRM structure. +++, Maximal activity equivalent to that of intact interacting proteins; –, no interaction; +/–, weak interaction; +, strong interaction; ++, nearly wild-type interaction.

Fig. 3. Expression of human *PUMILIO* mRNAs. *(A)* Northern analysis of adult tissues indicated PUM2 expression in ovary and testis. Two transcripts of approximately 5.5 and 6.5 kb were detected. *(B)* RT-PCR expression analysis of *PUM1* and *PUM2* in embryonic stem (ES) cells, fetal ovary and testis, adult brain and heart, and adult ovary and testis. *(C)* RT-PCR expression analysis of *PUM1* and *PUM2* in various fetal tissues and from men with normal germ cells (GCs) or no germ cells presence in their testis. Germ cell number varied between biopsies.

Fig. 4. PUM2 expression mirrored that of DAZ and DAZL. *(A)* Western analysis with PUM2 antisera (lanes 1-3), PUM2 preimmune antisera (lanes 4-6) and PUM2 antisera preincubated with PUM2 protein competitor (lanes 7-9). Cell extracts from human testis (lane 1,4,7), mouse testis (lane 2,5,8), mouse ES cells (lane 3,6,9). The major PUM2 protein band was ~ 95 kD in extracts from human testis, mouse testis and mouse ES cells.

Antisera specific to both DAZ and DAZL proteins (26) were used to stain human sections of *(B)* testis, *(C)* ovary, and *(D)* fetal testis. Antisera that recognized a PUM2 specific epitope were used to stain human sections of *(E)* testis (inset is a magnified cross section of stained tissue), *(F)* ovary, and *(G)* fetal testis. Preimmune sera were used to stain human sections of *(H)* adult testis, *(I)* adult ovary, and *(J)* fetal testis. Sections *(B-J)* counterstained with hematoxylin eosin; stained for specific proteins with horseradish peroxidase conjugated secondary antibodies (brown staining). Magnification: X200. Spg, spermatogonial cells; Spc, spermatocytes; Spd, spermatids; SC, Somatic cells; Oc, oocyte; RS, round spermatid.

Fig. 5. Expression of Pum2 in mouse. Pum2 expression was observed in *(A)* adult ovary, and *(B)* adult testis (inset is a magnified cross section of stained tissue). No staining was observed with PUM2 preimmune sera on *(C)* adult ovary or *(D)* testis. The pattern of PUM2 expression was also compared to that of mouse Boule. Boule protein was not present in *(E)* adult ovary and demonstrates a meiosis-specific expression pattern in *(F)* adult testis. Sections *(A-F)* counterstained with hematoxylin eosin; stained for specific proteins with horseradish peroxidase conjugated secondary antibodies (brown staining). Magnification: X200. Abbreviations as in Fig. 4.

Fig. 6. Interaction of PUM2, DAZL and NRE RNA. *(A)* Schematic of components used to detect PUM2/NRE interaction. Positive RNA-protein interaction was detected when the NRE and PUM2:AD were present as assayed by X-GAL assay (right panel). No interaction was detected with GAL4 activation domain (AD) or PUM2:AD without the

NRE or with NRE and AD alone. *(B)* Schematic of components used to detect NRE/PUM2/DAZL complex. A ternary complex was formed with NRE, PUM2, and DAZL:AD present (left panel). No interaction was detected with the NRE and the AD or DAZL:AD alone or with the NRE with PUM2 and the AD alone. No interaction was detected in the absence of the NRE, or with PUM2 and DAZL:AD present. Positive control interaction of the NRE with PUM2:AD is shown.

Table 1

Genes identified that encode proteins that potentially interact with DAZ

Gene	Chromosome location	No. of times obtained	Expression pattern	Motif
PUM2	2p22	2	Male/female germ cells and ES cells*	PUF repeat
hQK3	6q27	1	Upregulated in ovary and testis**	None
BOL	2q33	1	Male germ cells[39]	RRM, DAZ repeat
DZIP1	13q31.2	6	Specific tissues and testis**	Zinc finger
DZIP2	1p32.1	1	Male germ cells**	Zinc finger
DZIP3	3p13	2	Unknown	None
DAZL	3p24	1	Male/female germ cells[18,26,27]	RRM, DAZ repeat

GenBank accession nos.: PUM2, AF272350; hQK3, AF272349; BOL, AF272858; DZIP1, AF272347; DZIP2, AF272348; DZIP3, AF279370; DAZL, NM_001351.

*Expression demonstrated in this report.

**Expression determined by Moore and Reijo Pera (data not shown).

Abbreviations: ES, embryonic stem; RRM, RNA recognition motif.

References

1. Reijo, R., Lee, T. Y., Salo, P., Alagappan, R., Brown, L. G., Rosenberg, M., Rozen, S., Jaffe, T., Straus, D., Hovatta, O. & et al. (1995) *Nature Genet.* **10,** 383-93.

2. Reijo, R., Alagappan, R. K., Patrizio, P. & Page, D. C. (1996) *Lancet* **347,** 1290-93.

3. Simoni, M., Gromoll, J., Dworniczak, B., Rolf, C., Abshagen, K., Kamischke, A., Carani, C., Meschede, D., Behre, H. M., Horst, J. & Nieschlag, E. (1997) *Fertil. Steril.* **67,** 542-7.

4. Chai, N. N., Phillips, A., Fernandez, A. & Yen, P. (1997) *Mol. Hum. Reprod.* **3,** 705-708.

5. Carani, C., Gromoll, J., Brinkworth, M. H., Simoni, M., Weinbauer, G. F. & Nieschlag, E. (1997) *Mol. Hum. Reprod.* **3,** 479-483.

6. Gromoll, J., Weinbauer, G. F., Skaletsky, H., Schlatt, S., Rocchietti-March, M., Page, D. C. & Nieschlag, E. (1999) *Hum. Mol. Genet.* **8,** 2017-24.

7. Houston, D. W., Zhang, J., Maines, J. Z., Wasserman, S. A. & King, M. L. (1998) *Development* **125,** 171-80.

8. Johnson AD, B. R., Drum M, Masi T. (2001) *Dev. Biol.* **234,** 402-15.

9. Karashima, T., Sugimoto, A. & Yamamoto, M. (2000) *Development* **127,** 1069-79.

10. Maegawa, S., Yasuda, K. & Inoue, K. (1999) *Mech. Dev.* **81,** 223-26.

11. Eberhart, C. G., Maines, J. Z. & Wasserman, S. A. (1996) *Nature* **381,** 783-5.

12. Cooke, H. J., Lee, M., Kerr, S. & Ruggiu, M. (1996) *Hum. Mol. Genet.* **5,** 513-6.

13. Saxena, R., Brown, L. G., Hawkins, T., Alagappan, R. K., Skaletsky, H., Reeve, M. P., Reijo, R., Rozen, S., Dinulos, M. B., Disteche, C. M. & Page, D. C. (1996) *Nature Genet.* **14,** 292-9.

14. Seboun, E., Barbaux, S., Bourgeron, T., Nishi, S., Algonik, A., Egashira, M., Nikkawa, N., Bishop, C., Fellow, M., McElreavey, K. & Kasahara, M. (1997) *Genomics* **41,** 227-235.

15. Shan, Z., Hirschmann, P., Seebacher, T., Edelmann, A., Jauch, A., Morell, J., Urbitsch, P. & Vogt, P. H. (1996) *Hum. Mol. Genet.* **5,** 2005-11.

16. Yen, P. H., Chai, N. N. & Salido, E. C. (1996) *Hum. Mol. Genet.* **5,** 2013-7.

17. Houston, D. W. & King, M. L. (2000) *Development* **127,** 447-56.

18. Ruggiu, M., Speed, R., Taggart, M., McKay, S. J., Kilanowski, F., Saunders, P., Dorin, J. & Cooke, H. (1997) *Nature* **389,** 73-77.

19. Lin, H. & Spradling, A. C. (1997) *Development* **124,** 2463-76.

20. Forbes, A. & Lehmann, R. (1998) *Development* **125,** 679-690.

21. Asaoka, M., Sano, H., Obara, Y. & Kobayashi, S. (1998) *Mech. Dev.* **78,** 153-8.

22. Kraemer, B., Crittenden, S., Gallegos, M., Moulder, G., Barstead, R., Kimble, J. & Wickens, M. (1999) *Curr. Biol.* **9,** 1009-1018.

23. Parisi, M. & Lin, H. (1999) *Genetics* **153,** 235-50.

24. Crittenden, S. L., Bernstein, D., Bachorik, J., Thompson, B., Moulder, G., Barstead, R., Wickens, M. & Kimble, J. (2002) *Nature* **417,** 660-663.

25. Pryor, J. L., Kent-First, M., Muallem, A., Van Bergen, A. H., Nolten, W. E., Meisner, L. & Roberts, K. P. (1997) *N. Engl. J. Med.* **336,** 534-9.

26. Reijo, R. A., Dorfman, D. M., Slee, R., A. A. Renshaw, Loughlin, K. R., Cooke, H. & Page, D. C. (2000) *Biol. Reprod.* **63,** 1490-1496.

27. Dorfman, D. M., Genest, D. R. & Pera, R. A. R. (1999) *Hum. Reprod.* **14,** 2531-2536.

28. Neiderberger, C., Agulnik, A. I., Cho, Y., Lamb, D. & Bishop, C. E. (1997) *Mamm. Genome* **8,** 277-78.

29. Slee, R., Grimes, B., Speed, R. M., Taggart, M., Maguire, S. M., Ross, A., McGill, N. I., Saunders, P. T. K. & Cooke, H. J. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 8040-45.

30. McCarrey, J. R. (1993) in *Cell and Molecular Biology of the Testis*, eds. Desjardins, C. & Ewing, L. L. (Oxford University Press, New York), pp. 58-89.

31. SenGupta, D. J., Zhang, B., Kraemer, B., Pochart, P., Fields, S. & Wickens, M. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 8496-8501.

32. Tsui, S., Dai, T., Warren, S. T., Salido, E. C. & Yen, P. H. (2000) *Biol. Reprod.* **62,** 1655-60.

33. Venables, J. P., Ruggiu, M. & Cooke, H. J. (2001) *Nucl. Acids Res.* **29,** 2479-83.

34. Jiao, X., Trifillis, P. & Kiledjian, M. (2002) *Biol. Reprod.* **66,** 475-85.

35. Chubb, C. (1992) *J. Androl.* **13,** 312-7.

36. Ebersole, T. A., Chen, Q., Justice, M. J. & Artzt, K. (1996) *Nature Genet.* **12,** 260-5.

37. Ruggiu, M. & Cooke, H. (2000) *Gene* **252,** 119-126.

38. Wharton, R. P., Sonoda, J., Lee, T., Patterson, M. & Murata, Y. (1998) *Mol. Cell* **1,** 863-872.

39. Wang, X., Zamore, P. D. & Hall, T. M. T. (2001) *Mol. Cell* **7,** 855-65.

40. Edwards, T. A., Pyle, S. E., Wharton, R. P. & Aggarwal, A. K. (2001) *Cell* **105,** 281-9.

41. Xu, E. Y., Moore, F. L. & Pera, R. A. R. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 7414-19.

42. Sonoda, J. & Wharton, R. P. (2001) *Genes Dev.* **15,** 762-73.

43. Wickens, M., Bernstein, D. S., Kimble, J. & Parker, R. (2002) *Trends Genet.* **18,** 150-7.

Acknowledgements

We thank Michael Castillo for technical assistance, Anthony Dobson, Susan Fisher, Holly Ingraham, Richard Weiner, Joyce Tung and Eugene Xu for comments on the manuscript. We also thank Marvin Wickens for providing us with three-hybrid resources. This work was supported by the Ford Foundation and Woodrow Wilson Foundation fellowships to F.L.M., by the Polish State Committee for Scientific Research (J.J.), by an NIH-NRSA fellowship (J.U.), by a Lalor Foundation Grant (M.F.) and by grants from the National Institutes of Health (Ira Herskowitz and R.A.R.P), the Searle Foundation and the Sandler Family Foundation (R.A.R.P.).

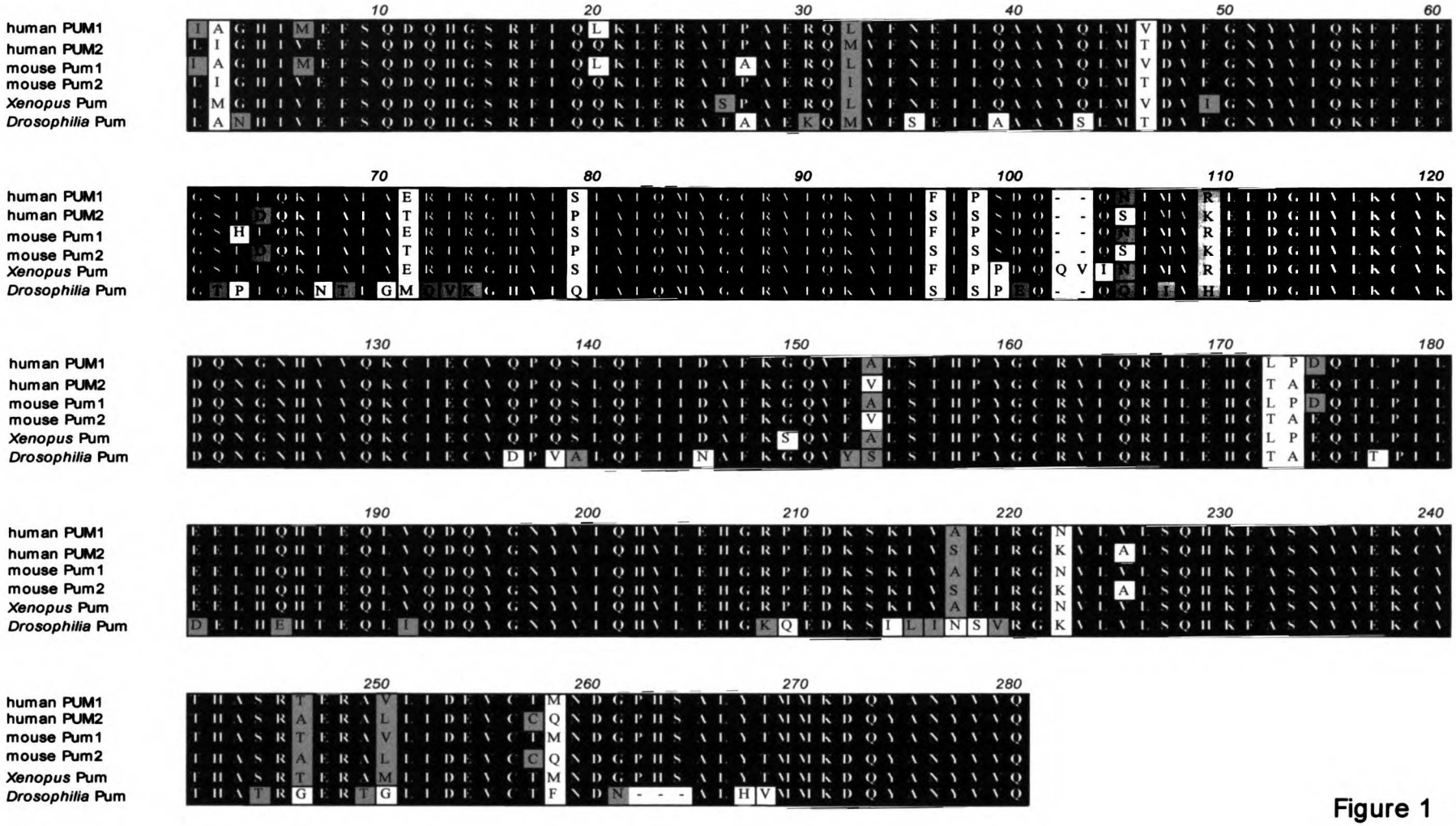

human PUM1
human PUM2
mouse Pum1
mouse Pum2
Xenopus Pum
Drosophilia Pum
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210
220
230
240
250
260
270
280

Figure 1

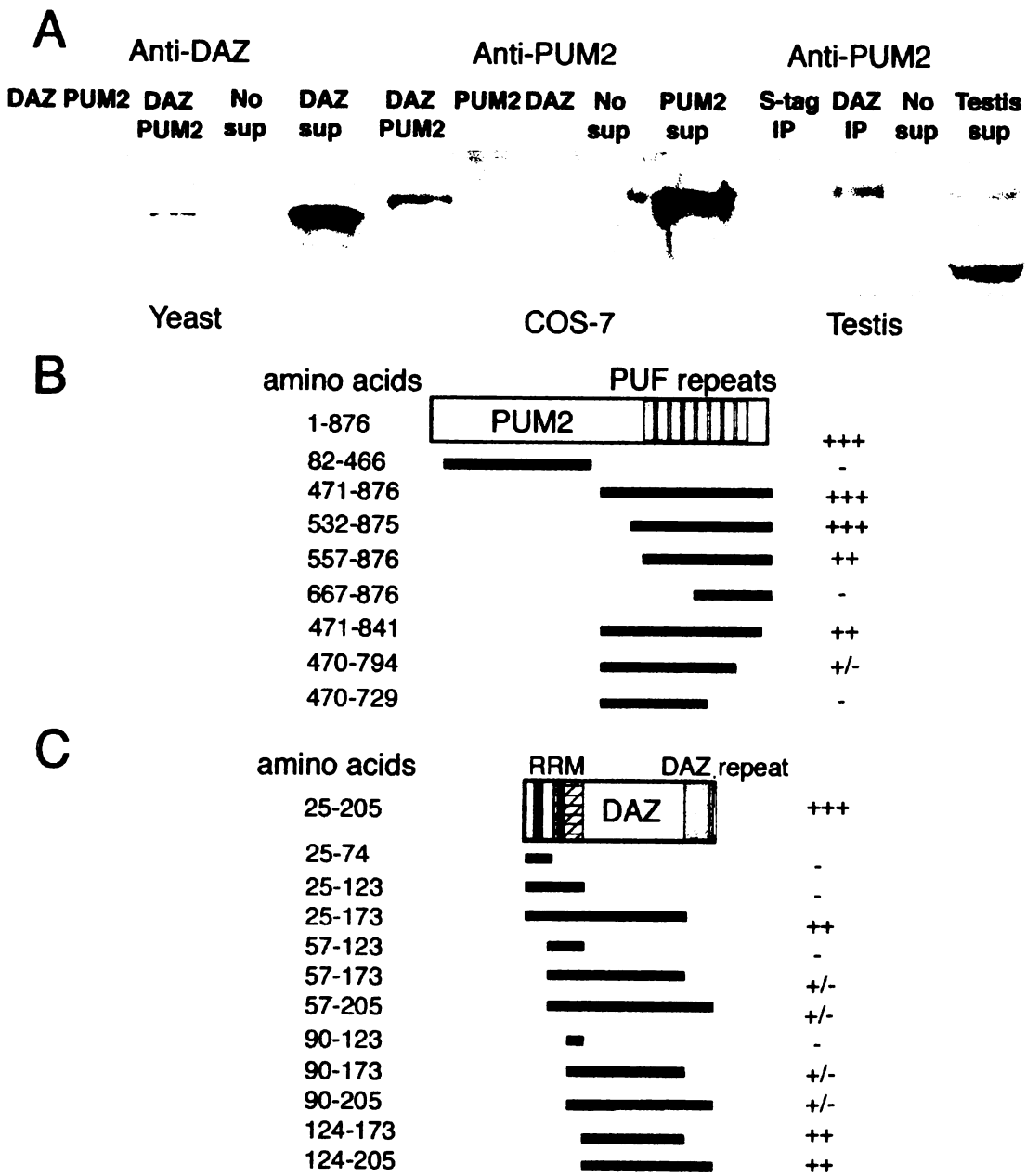
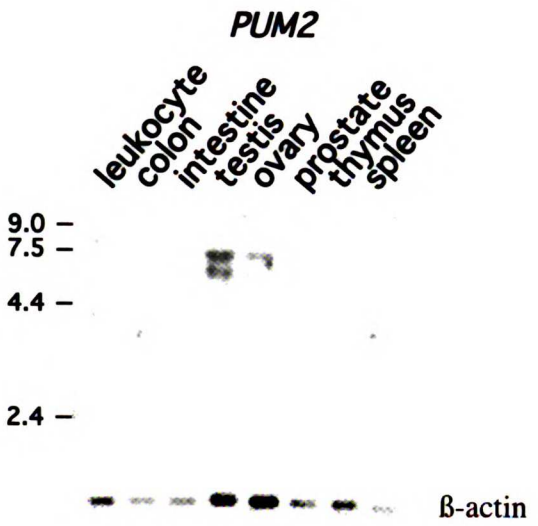
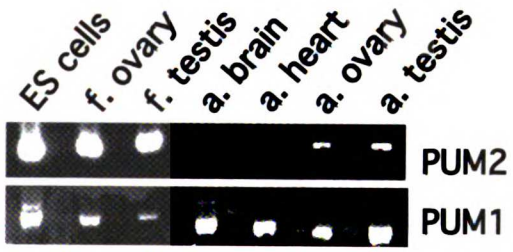

A
Anti-DAZ
Anti-PUM2
Anti-PUM2
DAZ PUM2
DAZ PUM2
No sup
DAZ sup
DAZ PUM2
PUM2 DAZ
No sup
PUM2 sup
S-tag IP
DAZ IP
No sup
Testis sup
Yeast
COS-7
Testis
B
amino acids
PUF repeats
PUM2
1-876 +++
82-466 -
471-876 +++
532-875 +++
557-876 ++
667-876 -
471-841 ++
470-794 +/-
470-729 -
C
amino acids
RRM
DAZ repeat
DAZ
25-205 +++
25-74 -
25-123 -
25-173 ++
57-123 -
57-173 +/-
57-205 +/-
90-123 -
90-173 +/-
90-205 +/-
124-173 ++
124-205 ++

Figure 2

A


PUM2
leukocyte
colon
intestine
testis
ovary
prostate
thymus
spleen
9.0
7.5
4.4
2.4
ß-actin

B


ES cells
f. ovary
f. testis
a. brain
a. heart
a. ovary
a. testis
PUM2
PUM1

C

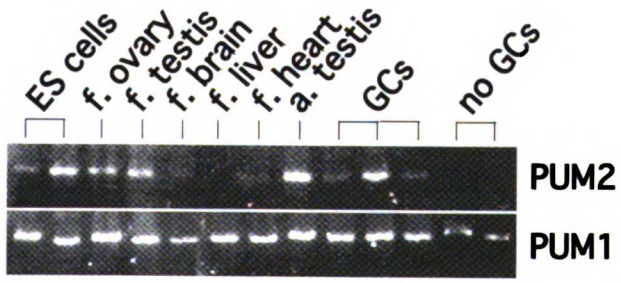

ES cells
f. ovary
f. testis
f. brain
f. liver
f. heart
a. testis
GCs
no GCs
PUM2
PUM1

Figure 3

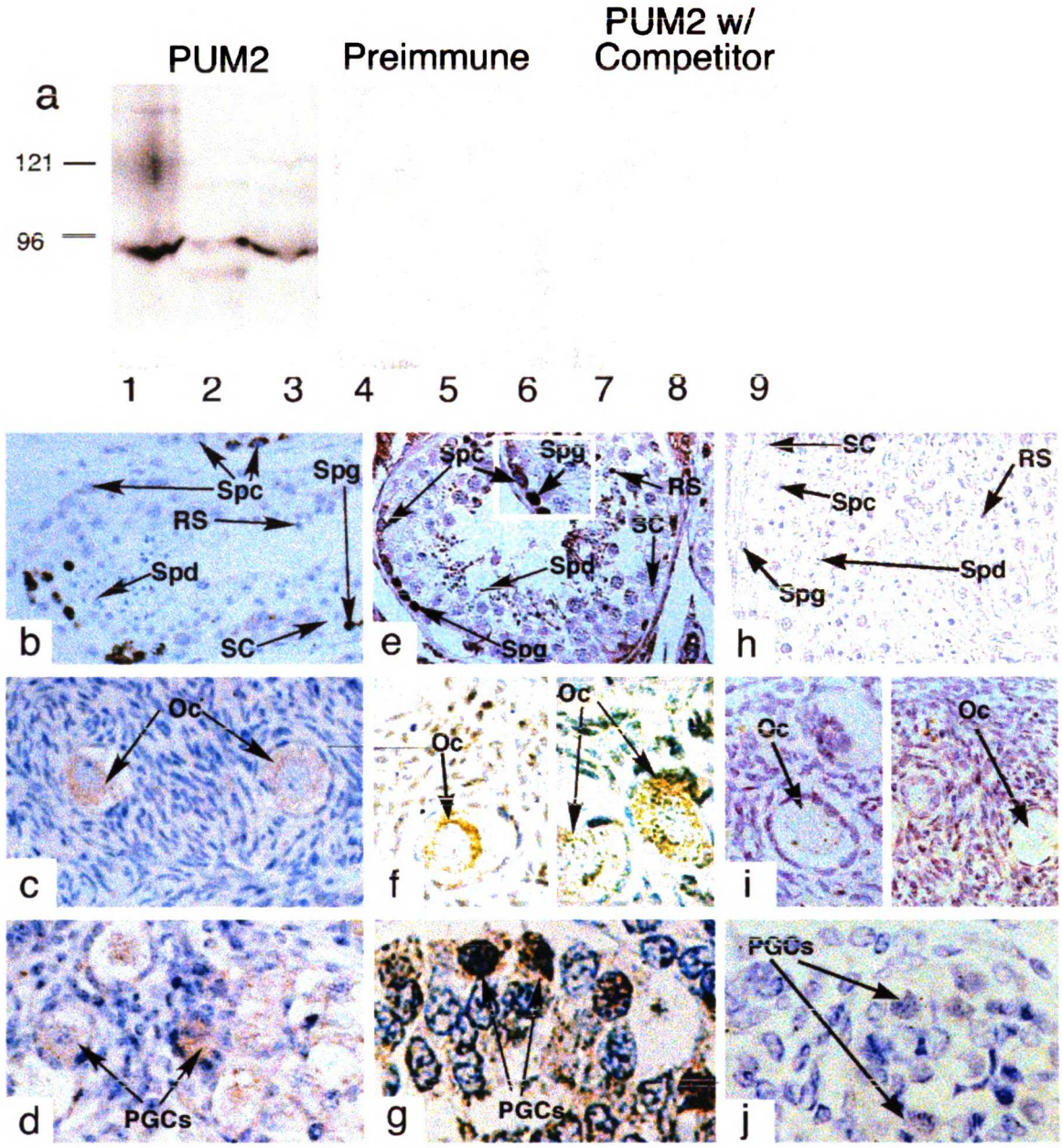

PUM2
Preimmune
PUM2 w/
Competitor
a
121
96
1 2 3 4 5 6 7 8 9
Spc
Spg
RS
Spd
SC
b
e
h
Oc
c
f
i
PGCs
d
g
j

Figure 4

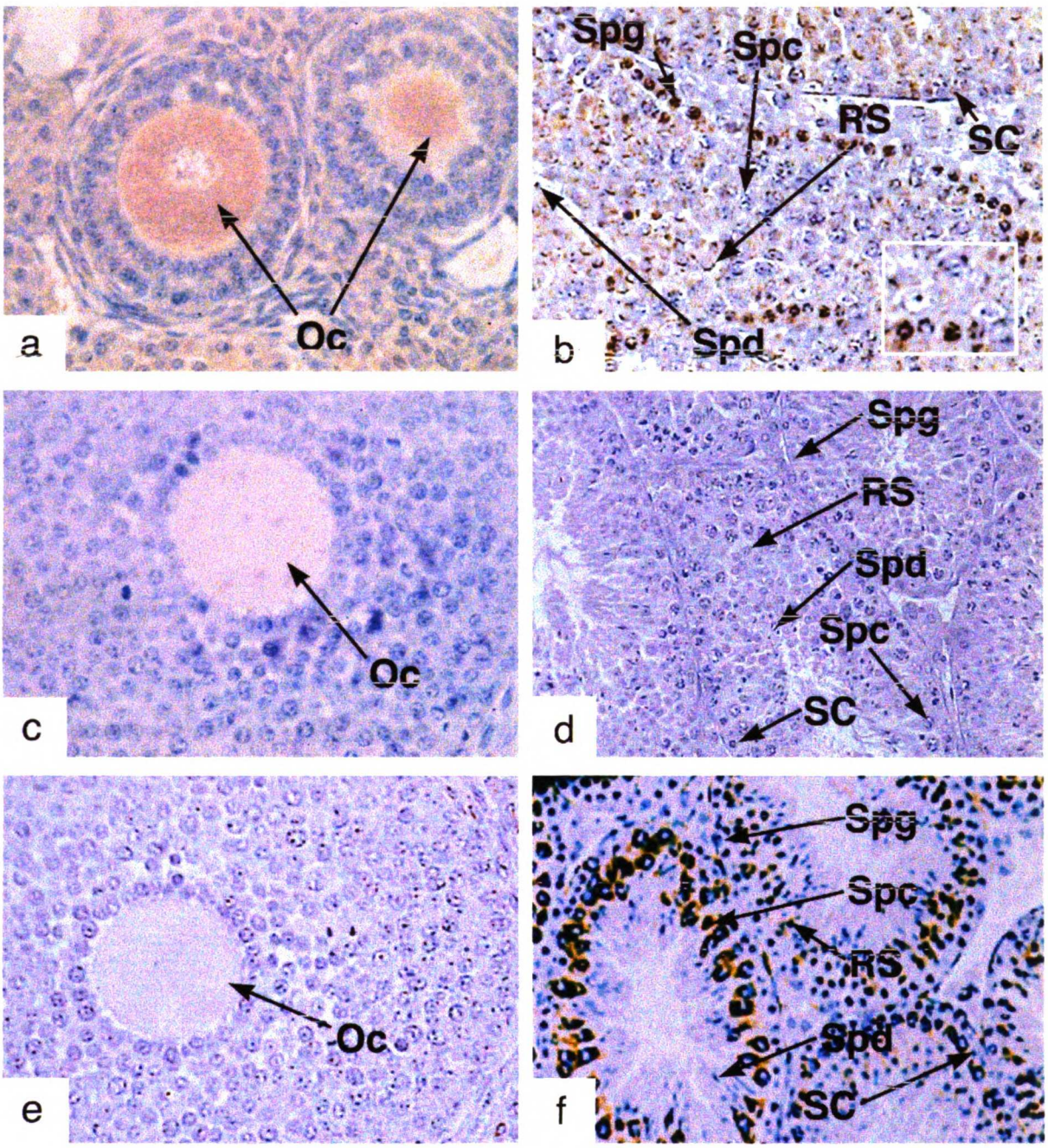
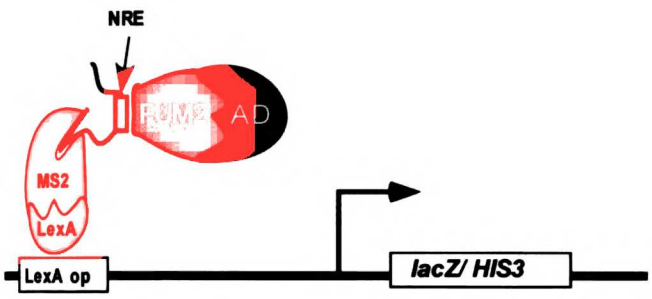


Figure 5

a



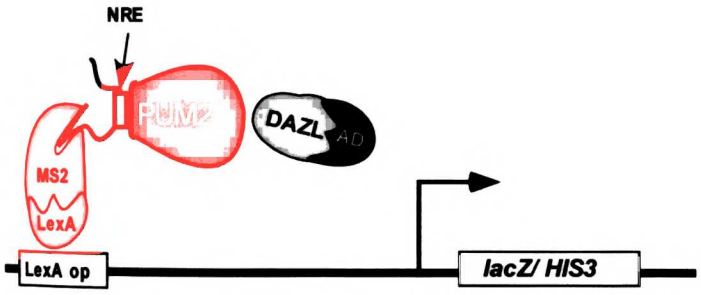
AD

PUM2:AD

NRE, AD

NRE, PUM2:AD

b



NRE, AD

NRE, DAZL:AD

NRE, PUM2, AD

NRE, PUM2, DAZL:AD

PUM2, DAZL:AD

NRE, PUM2:AD

Figure 6

Chapter 3

A gene family required for human germ cell development evolved from an ancient meiotic gene conserved in all metazoans

A gene family required for human germ cell development evolved from an ancient meiotic gene conserved in all metazoans

Eugene Yujun Xu, Frederick L. Moore and Renee A. Reijo Pera*

Department of Obstetrics, Gynecology and Reproductive Sciences; Department of Physiology; Department of Urology; and Programs in Human Genetics and Cancer Genetics; University of California at San Francisco, San Francisco, CA 94143-0546

Abstract

Two paralogs of the *Deleted in AZoospermia (DAZ)* gene family exist in humans: *DAZ,* a Y-chromosome gene whose deletion causes male infertility due to production of few or no sperm (1,2) and *DAZL (DAZ-Like)*, an autosomal gene required for premeiotic germ cell development in both females and males (3-6). Homologues of *DAZ* have also been identified in diverse species yet their function ranges from meiotic regulation to primordial germ cell development (6-12). Here we report the identification of a new member of the *DAZ* gene family, *BOL*, which was identified in a two-hybrid screen for DAZ interacting proteins. Both human and mouse *BOL* resemble the invertebrate meiotic regulator, *boule*, the proposed orthologue of *DAZ* (13), in sequence and expression pattern and hence likely perform a similar meiotic function. In contrast, human *DAZ* and *DAZL* are expressed much earlier in prenatal germ stem cells and in spermatogonia; *DAZL* is also expressed in female germ cells. Our data suggest that homologues of the *DAZ/BOL* gene family can be grouped into two subfamilies-*BOL* and *DAZL*, and that members of the human DAZ family evolved from an ancestral meiotic regulator, *boule,* to assume distinct, yet overlapping, functions in human germ cell development. We further verify the interaction between BOL and DAZ by coimmunoprecipitation in mammalian cells and identify the specific region of DAZ protein that BOL binds. These data suggest a conserved functional role for BOL from metazoans to humans and suggests a new acquired role for DAZ and DAZL proteins early in germ cell development.

Introduction

Many reproductive genes, especially in males, evolve rapidly in diverse lineages of metazoan animals including human, possibly driven by natural selection (14-16). Thus, it is paradoxical that major steps of germline development such as germ cell allocation, primordial germ cell migration and differentiation, mitotic proliferation and meiosis are maintained throughout metazoan evolution while the genes evolve at a fast rate. In our investigation of the human reproductive gene family, *Deleted in AZoospermia (DAZ)*, we uncovered a new family member that is conserved from flies to humans. We have reconstructed an evolutionary history of this germline-specific family that enables us to draw a picture that illustrates how the conservation of cellular steps during germ line development has occurred by conserving a critical genetic module in invertebrates and vertebrates dedicated to meiosis and modifying a copy of this module to accomplish a novel function in the vertebrate lineage.

The *DAZ* genes were identified in a screen for genes that map to the Y chromosome and are required for fertility in men (1). Deletions encompassing the Y chromosomal *DAZ* genes are the most common molecularly-defined cause of infertility in humans (17-19). An array of 4 *DAZ* genes arranged in two clusters are located on the Y chromosome and encode RNA-binding proteins with a common RNA-recognition motif (RRM) and a series of 8 - 18 consensus *DAZ* repeats, a carboxyterminal domain of 24 amino acids each that are rich in N, Y, and Q residues (1, 8, 20). An autosomal homolog of *DAZ, DAZL (DAZ-Like)* also maps to chromosome 3 (4,5,21-23). The predicted protein product of the

DAZL gene is >95% identical to that of *DAZ* except that *DAZL* contains just one *DAZ* repeat.

Homologs of *DAZ* have been identified in diverse organisms from invertebrates to vertebrates (3,6,10,24,25). The homologs in all these organisms are required for germ cell development but differ in null phenotypes and expression patterns. In flies, disruption of the *DAZ* homolog, *boule,* causes meiotic arrest during spermatogenesis (7). In *C. elegans,* disruption of the homolog of *DAZ* causes meiotic arrest in oogenesis but not spermatogenesis (10), and in mice, disruption of the *DAZ* homolog *Dazl* causes loss of all germ cells prenatally in both males and females (6). Furthermore, in zebrafish and in frogs, *DAZ* homologs are components of the germ plasm, a region of the oocyte cytoplasm that allocates the germ lineage and is composed of clusters of RNAs, RNA-binding proteins, ribosomes and mitochondria that segregate to give rise to germ cells (8-12). In frogs, *Dazl* has been shown to be required for germ cell production and migration during embryogenesis (9). Thus in diverse organisms, *DAZ* homologs are required for germ cell development, but differ in null phenotypes. Below we report the identification and characterization of a new member of the human *DAZ* gene family, *BOULE.* With the identification of *BOULE,* our phylogenetic analyses indicate that the *DAZ* gene family is composed of two subfamilies required for different stages of germ cell development: *DAZL* for early germ cell function and *BOL* for meiotic function. *BOL* is the ancestral gene that is conserved from flies to humans, whereas, *DAZL* arose in the early vertebrate lineage and *DAZ* arrived on Y chromosome during primate evolution (4).

Materials and Methods

Two-hybrid screening of DAZ interacting proteins

We used the yeast two hybrid system to screen for proteins that interact with DAZ protein expressed as a fusion with the GAL4 DNA binding domain in a pAS2 vector (Clontech, Palo Alto, CA). The DAZ construct was derived from a Y chromosome-encoded cDNA, pRR102, which encodes the N-terminal RNA-binding domain and a single *DAZ* repeat. The protein expressed by this plasmid is DAZ and is 98% identical to that of DAZL. A testis library of random cDNAs fused to the GAL4 activation domain was transformed into yeast containing the DAZ fusion; $9X10^6$ yeast transformants were screened for interaction with DAZ according to manufacturer's instructions (Clontech, Palo Alto, CA).

Sequence analyses

Human *BOULE* sequence was obtained by sequencing a clone we obtained from two-hybrid analysis that contained the entire open reading frame (Human Genetics Core Facility, UCSF). Mouse *Boule* was assembled from an EST clone, accession Number: AI789678. A mouse testis cDNA was cloned from a mouse testis cDNA library (Genome Systems, St. Louis, MO). ClustalW (1.4) alignment of MacVector 6.5 was used for multiple sequence alignments; parameters were open gap penalty of 10, extend gap penalty of 0.1, delay divergence of 40%, gap distance of 8, and a similar matrix as Blossum.

Western blotting and Immunocytochemistry

Polyclonal antibodies were raised in rabbits by injection of a synthetic oligopeptide (ETQEDAQKILQEAEKLNYKDKKLN) coupled to a MAP (Multiple-Antigen Peptide; Research Genetics, Inc., Huntsville, AL). Twelve to twenty weeks after initial injection of peptides, serum antibodies were purified as described (22,26). Testes tissue extracts were used for Western blotting as described and Western blotting was done with anti-Boule antisera (1:1000) also as described (22,26). Mouse testis sections were from 60-day adult mice. Human testis sections were from an adult male with complete spermatogenesis but physical obstruction of ducts. Images were captured through a Leica Microscope and assembled in Photoshop 5.5.

Northern and radiation hybrid mapping

Northern hybridization was done on polyadenylated RNA blots of human and mouse tissues (Clontech, Palo Alto, CA) using human and mouse cDNA clones. For radiation hybrid mapping, primers (38INSF and Hb18: GAGGAGGTGGATGTGACCC and CTTATTGCTGGACCAATGTTCA) were used to amplify a 1.5 kb fragment encompassing the second intron of BOL that is present in humans but not in hamsters. A second primer set (TTTTTCCATTCAGTCTTCCTGA and GCAGAGAAAATAAAAGACACCTCA) was used to confirm linkage. PCR of radiation hybrid panels GeneBridge 4.0 (Research Genetics, Inc.) was done on duplicates for 38INSF-Hb18 and once for exon 11 primers. For mouse mapping, the primers GCTCAGTTGCAGTGTGTTTTTC and TGCTCCATTCCTATATCTGCAA) were used and the consensus of two duplicate experiments was used to map the position.

BOL and DAZ interaction and deletion analysis

Coimmunoprecipitation of BOL and DAZ proteins from COS-7 cells was accomplished by transfection with a pBud vector (Invitrogen, Inc., Carlsbad, CA) containing the *BOL* cDNA fused to an HA tag (Roche, Inc., Indianapolis, IN) expressed from an EF-1 alpha promoter, and a pBud vector containing the *DAZ* cDNA expressed from a CMV promoter. Cells were cultured for ~48 hours, lysed with lysis buffer (50 mM Tris-HCl, pH 7.5, 1 mM EDTA, 150 mM NaCl, 10% glycerol, 1% NP-40) and incubated with HA antibodies covalently linked to inert beads (BABCO, Inc., Berkeley, CA) for two hours at 4 degrees. HA-beads were washed with lysis buffer 3 times, and samples were analyzed by Western analysis with antibodies to the DAZ protein. To construct truncated DAZ protein, primers were used to clone fragments of the *DAZ* cDNA in frame with the GAL4 DNA-binding domain within the pAS2 vector (Clontech Inc.). The constructs were transformed into Y190 yeast strain along with the BOL:GAL4 activation domain fusion construct and screened for protein-protein interactions by X-Gal filter assays (Clontech, Inc.).

Results and Discussion

Identification of human *BOULE*.

In order to better understand the molecular function of the human *DAZ* gene family, we sought to identify proteins that interact with DAZ/DAZL proteins. Using a DAZ/DAZL construct as bait in a two-hybrid screen, we identified a gene remarkably similar to *DAZ* and *DAZL*. However, the protein encoded by this gene is more homologous to *Drosophila* Boule, the proposed fly orthologue of DAZ (13), than to either DAZ or DAZL.

Protein sequence alignments of fly Boule, human DAZ, DAZL, and the product of this new gene identified in our screen called *BOL* show that BOL is a closer homolog to *Drosophila* Boule than either human DAZ or DAZL. The overall identity and similarity between human BOL and *Drosophila* Boule is greater (30% identity, 42% similarity) than that between either DAZ or DAZL and *Drosophila* Boule (13% identity and 19% similarity or 22% identity and 32% similarity respectively). Like *DAZL*, *BOL* encodes an RNA-binding protein that contains a single RNA binding domain with signature RNP-1 and RNP-2 motifs (Fig. 1*A-B*). The extensive homology shared by the RNA binding domains of fly Boule, human BOL, DAZ and DAZL distinguishes these proteins as a unique family of RNP proteins, the DAZ/BOL family. There is 80% similarity between the domains of BOL and Boule, 59% between DAZ and Boule, and 61% between DAZL and Boule. We note that BOL and *Drosophila* Boule share 21 amino acids between them that differ in DAZ while, in contrast, only 13 amino acids are shared between DAZ and

Boule that differ in BOL (Fig. 1*A*). Furthermore, BOL and Boule have identical RNP-1 and RNP-2 motifs while DAZ and DAZL differ slightly. A DAZ repeat is also present in BOL, though the repeats are less conserved than the RNA-binding domains (Fig. 1*C*).

Human BOULE Maps to Chromosome 2

To establish that human *BOL* is not simply a variant copy of the Y chromosome *DAZ* gene cluster, we mapped the *BOL* gene by radiation hybrid mapping and placed *BOL* on chromosome 2 at a position 0.4 cR (centiRads) from the marker D2S348 (LOD = 21). This mapping places *BOL* in human 2q33, a position verified by mapping of two genomic BAC clones by fluorescence in-situ hybridization (FISH; data not shown). We further verified these findings by isolating a cDNA fragment from mouse testes and mapping mouse *Boule* to its chromosomal position. The predicted protein of a mouse cDNA clone is 86% identical to human BOL and its position on chromosome 1 is syntenic to human 2q33.

Expression of DAZ/DAZL and BOL proteins

We next compared the expression patterns of the different members of the DAZ/BOL family using antibodies specific to DAZ, DAZL and BOL. Expression of the DAZ and DAZL genes is known to be restricted to germ cells (3,5,6,8,12,22,26-30). As shown (Fig. 2 *A* and (22,26)*), DAZ* and *DAZL* are expressed early in development in both the human male and female. Staining of the nucleus and cytoplasm of primordial germ cells is abundant in both sexes. In the adult, both proteins are expressed only in germ cells with expression being most abundant in the nucleus and cytoplasm of spermatogonia and

in the cytoplasm of the meiotic spermatocytes (Fig. 2*B-C)*. DAZL is also abundant in the cytoplasm of oocytes, as well (22). The cellular and subcellular expression pattern of mouse Dazl is identical to that of the human (Fig. 2*D)*. Thus, consistent with the phenotypes observed when vertebrate DAZL homologs are disrupted, the expression of the DAZL and DAZ genes begins early in development in the germ stem cell populations and continues through the meiotic divisions of gametogenesis.

Considering that the human and mouse *BOL* genes are members of the *DAZ* gene family, we next investigated whether the genes are transcribed and whether the proteins are found in the same cellular and subcellular compartments as the DAZ and DAZL proteins. When we examined expression of the *BOL* genes in both mice and men, we found that in both species, *BOL* transcripts are restricted to the testes (Fig. 2*E-F)*. Western blotting with antisera specific to BOL also shows that a 32 KD BOULE protein is confined to testes in both human and mouse (Fig. 2*G*). Thus, superficially, the expression of BOULE resembles that of human DAZ and DAZL.

However, when we examined the cellular and subcellular distribution of BOL in mouse and human testes, results indicated that BOL is not present in spermatogonial cells but instead is first found in detectable quantities in the cytoplasm of spermatocytes and then persists through meiosis in both species (Fig. 2*H-K*). In mice, which are amenable to staging of germ cell types, BOL expression begins in stage III spermatocytes and peaks in late pachytene or diplotene stage spermatocytes (Fig. 2*I-K;* low power magnification is shown in Fig. 2*L-M*). BOL is detectable in secondary spermatocytes and

early spermatids, then gradually decreases until it is undetectable in late spermatids. Thus, surprisingly the expression of BOL is distinctly different from that of DAZ and DAZL and instead the expression of human BOL is identical to that of fly *boule* (13,31).

In light of our observations on analysis of sequences and expressions of the human DAZ/BOL family, we propose that human *BOL* is the orthologue of *Drosophila boule; DAZ/DAZL*, on the other hand, are paralogues of *BOL*. This hypothesis is supported by the protein sequence comparison and the null phenotypes of known *DAZ* and *Boule* homologs.

Analysis of BOL interaction with DAZ

We next sought to verify that the DAZ protein interacts with BOL protein by co-immunoprecipitation in mammalian cells. Supernatant was derived from COS-7 cells expressing DAZ alone, BOL alone, and both DAZ and BOL. We utilized a HA-tag, which is fused to the N-terminus of BOL to isolate BOL proteins from mammalian supernatant and probed using Western analysis for the presence of DAZ protein. Our results show that the DAZ protein is only isolated from the mammalian supernatant containing both DAZ and BOL proteins (Fig. 3*A*). These data shows that the DAZ protein can form a stable complex with BOL protein. To define the minimal regions necessary for DAZ and BOL interactions, we used deletion analysis with various constructs in the yeast two-hybrid system. Analysis of cDNA deletions that produce truncated DAZ protein indicated that the minimal region of DAZ necessary for interaction with BOL is the linker region of DAZ protein between the RNA-binding

domain and the DAZ repeat region (amino acids 124-173 contain the minimal DAZ binding domain; Fig. 3*B*).

The DAZ/BOL Family Contains Two Subgroups

Alignments of the known DAZ homologs and their RNA binding domains from major groups of vertebrates and invertebrates suggest that the DAZ/BOL family can be divided into two subgroups—*DAZ* and *BOL*. The distance tree of DAZ/BOL proteins shows that homologs in zebrafish and frog are closer to DAZL in overall sequence and worm *Daz-1* is almost equally divergent from *BOL* and *DAZ* but is slightly closer to the *BOULE* group (Fig. 4*A*). This suggests that the homologs in zebrafish and frog are indeed *DAZL* homologs while worm *Daz-l* is a BOL homolog. This is corroborated by distinct features in the RNA binding domains unique to each subgroup of *DAZ/BOL* family (Fig. 4*B*). First there is a three amino acid deletion in position 74 to 76 shared by human, mouse, frog and zebrafish *DAZ* homologs that is not present in human *BOL(BOL.hs)*, fly *Boule(Bol.dm)* and worm *Daz-1(cDazl)*. Second, in the most conserved RNP-1 and RNP-2 motifs, all *Boule* homologs encode amino acids IFVGG while *DAZ* homologs encode V/LFVGG. RNP-1 in *Boule* encodes KGYGFV/IT, but KGYGFV/IS/Y in *DAZ*. The sequence at the C-terminus of the RNA binding domain yields the same grouping.

The expression and null phenotypes of *DAZ* and *boule* homologs further support the existence of two subgroups within DAZ/BOL family required for distinct functions—BOL for meiotic function and DAZL for primordial germ cell function (Table

1). *Boule* is required for entry into and progression through meiosis in *Drosophila* males. The *C. elegans BOL* homolog is required for female meiosis; oocytes are arrested prematurely in meiosis in null worms (10). In contrast, mouse *DAZ* homolog null mutants have prenatal, premeiotic germline defects in both the male and female (6). In zebrafish and frogs, *DAZ* homologs encode components of the germ plasm, a region of the oocyte cytoplasm rich in clusters of RNAs, RNA-binding proteins, ribosomes and mitochondria that gives rise to germ cells. Recently Houston and King showed that the *Xenopus DAZ* homolog is localized in germ plasm and early precursor germ cells and is required for proliferation or differentiation of germ stem cells (8,9). Such different expressions and functions among DAZ/BOL family members can be explained by proposing that *BOL*, not *DAZL*, is the orthologue of *Drosophila boule* and that homologs of DAZL and BOL perform functions at different stages of germ cell development.

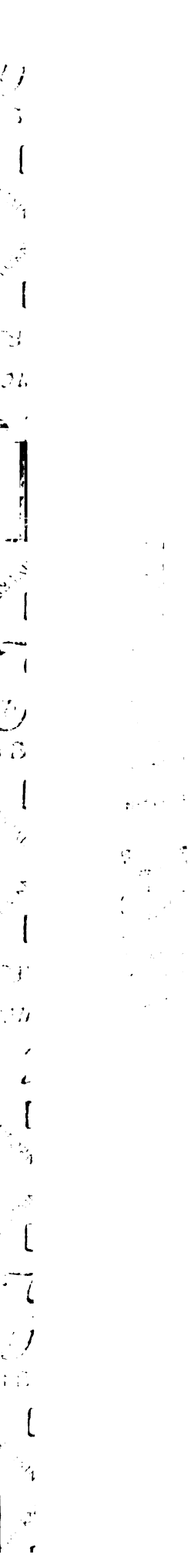
Boule is the DAZ/BOL Ancestor of Invertebrate and Vertebrate Lineages

We next explored the evolutionary history of DAZ/BOL family. It is certain that only one *BOL* homolog exists in flies and worms and there are no *DAZ* homologs in these organisms since both genomes have been completely sequenced, yet both *BOL* and *DAZL* exist in mammals. Thus, *DAZL* was either lost in the invertebrate lineage or arose during vertebrate evolution. Comparison of genomic structure among human *DAZL* genes, human *BOL* and fly *boule* supports a rise of *Dazl* from *Bol* during early vertebrate lineage. We determined the genomic structure of *BOL* and found that, like *DAZL* and *DAZ*, *BOL* has 11 exons and shares 6 splicing positions with *DAZ/DAZL* (3 in the RNP domain, Fig. 1*b*). Eight splice positions are shared if we consider junctions with a single

deletion/addition difference. The observations that there are more exon-intron splicing sites shared between human *BOL* and *DAZL* than between *BOL* and fly *Bol*, and that there is an identical number of exons in BOL and DAZL, suggest the historical relationship between BOL and DAZL is closer than that between BOL and Boule. Such observations argue against the loss of DAZL in invertebrates. This is not in conflict with the higher homology of protein sequence between BOL and Bol, which reveals a similar functional constraint upon these proteins but not necessarily a closer ancestral relationship. Hence, we propose that the single copy gene, *Boule,* is the *DAZ/BOL* ancestor of the vertebrate and invertebrate lineages. *DAZL* arose later in the vertebrate lineage through gene duplication. Based on sequence alignments and expression patterns of the *DAZL* homologs in frog and zebrafish(8,9,11,12), DAZL arose through a duplication of *Boule* that occurred before the divergence of modern vertebrates, but after splitting of invertebrates and vertebrates (Fig. 4*C*). Later during early primate evolution, after splitting of new world monkey (NWM) and old world monkey, another duplication took place from *Dazl* resulting in a new Y-linked *DAZ.* The Y-linked *DAZ* went through two more duplications as recent as 55,000 years ago giving rise to a cluster of four *DAZ* genes (4,19,35; Fig. 4*C*).

The Newly Emerging *DAZL* Gene Acquired a Novel Function in Vertebrates

After duplication of *Boule,* the newly emerging *DAZL* gene acquired a novel function in prenatal germline development while *BOULE* maintained its traditional role in meiosis. This is evident by the premeiotic expression in germ cells of both sexes and an essential requirement in germ stem cell development in the frog, mouse, and humans



(6,9, 22, 26). In frogs and mice, disruption of *DAZL* homologs leads to loss of premeiotic stem cell populations early in development, whereas, in contrast in flies and worms, disruption of the *BOL* homologs leads to meiotic defects late in germ cell development. Remnants of the traditional role for *DAZL* being maintained in vertebrates is evident by the observation that *Xenopus Dazl* can rescue a fly *boule* meiotic mutation (8). This suggests that although the *DAZL* genes have acquired a new function, they preserve an RNA binding domain that can bind the same RNA as *Boule* and can substitute for *Boule* when expressed in the right place at the right time. The novel function of DAZL has not come at the expense of the functional unit of Boule. Instead the Boule RNA binding domain has been used as a module for evolutionary innovation.

Although germ cells provide continuity of life from generation to generation, evidence indicates that many genes required for germ cell function, especially in males, evolve rapidly in species as diverse as *Drosophila,* marine invertebrates, and primates including humans (15,16). Here we identified a new member of DAZ gene family, *BOL*, which is conserved from flies to humans. We further traced the evolution of this human reproductive gene family from an ancient meiotic regulator to a family of genes with different, yet overlapping, functions in germline development (Fig. 5). Mammalian BOULE resembles the invertebrate meiotic regulator, Boule, in sequence and expression pattern and likely performs a similar meiotic function in the meiotic cells in which it is expressed. In contrast, DAZL is present only in vertebrates and is required for prenatal primordial germ cell function, indicating development of a novel feature required for genetic control of vertebrate germline development but not invertebrate germline

development. Finally, in contrast to both *DAZL* and *BOULE*, *DAZ*, the youngest member of the family, is not essential for completion of spermatogenesis as evidenced by two facts. First, deletions encompassing the Y chromosome cluster cause phenotypes that vary from complete absence to low number of sperm in the semen; and second, human DAZ exons and introns are evolving at the same rate, indicative of neutral genetic drift (1,2,32). Perhaps *DAZ*, although implicated in quantitative aspects of sperm production, has yet to evolve a function essential for completion of spermatogenesis. The distinct, yet overlapping, expression patterns of the human DAZ/BOL family suggests that through evolution, different members of the family have gradually acquired new functions to eventually encompass the major framework of the human reproductive pathway (Fig. 5) and that perhaps, the *DAZ* genes, which duplicated twice in recent history (20), are probing the evolutionary possibilities of nature. Whether the emergence of the *DAZ* gene cluster on the human Y chromosome is a reinforcement of the ancient meiotic function of *Boule* or a novel function in the making remains to be seen as *DAZ* genes continue to evolve.

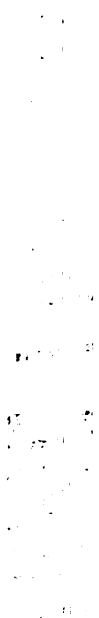


Figure Legends

Figure 1. Human *BOULE (BOL)* encodes a protein that is homologous to fly Boule and human DAZ/DAZL. (A) Diagram of fly Boule (Bol.dm) and, human BOL, DAZL and DAZ proteins. Black boxes represent the RNA binding domains, hatched boxes represent the DAZ repeats. (B) Alignment of RNA binding domains of fly Boule, human BOL, human DAZL and human DAZ proteins. (*) indicates conserved residues and (.) similar residues among all four proteins. Shadowed box outlines amino acids shared between Boule and BOL but different from DAZ or DAZL. Open box indicates amino acids shared between DAZ or DAZL and BOL but not with fly Boule. Solid arrows indicate shared splice sites and open arrows indicate unique splice sites in Boule and/or BOL. Dashed arrow indicates a shared splice site if we consider the two amino acids (position#73-#74) as a single addition/deletion. (C) Conservation of the DAZ repeats in DAZ and BOL homologues. Shade indicates identical or similar amino acids.

Figure 2. BOL expression is distinct from that of DAZ and DAZL. DAZ is expressed in prenatal primordial germ cells, spermatogonial stem cells and spermatocytes. DAZL is expressed in both the male and female germline. BOL is expressed in the cytoplasm of early spermatocytes, persists through meiosis, and decreases in early spermatids. (A) DAZ staining of fetal primordial germ cells of the human testes; similar staining is observed in female primordial germ cells (22,26). (B) Human testes staining with antisera that recognize DAZL only; DAZL is expressed in spermatogonia, early and late spermatocytes, and postmeiotic cells. Staining of human ovary with this antisera indicate

cytoplasmic staining of oocytes (22,26). (C) Human testes section stained using antisera that recognize DAZ only. DAZ is expressed in spermatogonia and early spermatocytes, but is absent from late spermatocytes or postmeiotic cells. (D) Mouse Dazl is present in spermatogonia and early and late spermatocytes as in human testes. As in humans, female mice also express Dazl in the germ cells (6,22). (E) A northern blot with polyadenylated RNA from different human tissues. Blot was hybridized with human BOL cDNA that detects two testes specific transcripts. 1-spleen, 2-thymus, 3-prostate, 4-testis, 5-ovary, 6-small intestine, 7-colon, 8-leukocyte. (F) A northern blot with polyadenylated RNA from mouse tissues. Blot was hybridized with mouse BOL cDNA that detects three testes specific transcripts. 1-heart, 2-brain, 3-spleen, 4-lung, 5-liver, 6-muscle, 7-kidney, 8-testis. (G) Anti-BOULE antisera detects a single 32 KD protein in mouse testes (b) and a similar size protein in human testes (a) but not in other human or mouse tissues (data not shown), nor does it recognize DAZ protein expressed in yeast strain, RRY618 (c). The 50Kd band in human testes is non-specific as it is detected by preimmune also. (H) BOL staining in human testis section is also restricted to cytoplasm of spermatocytes; no staining of spermatogonial stem cells is observed. (I) Stage III seminferous tubules of mouse testis. BOL is expressed in round spermatids (Spd) and secondary spermatocytes but not in spermatogonia (Spg) or primary spermatocytes (Spc). (J) Stage VII seminiferous tubule of mouse testis. BOL is expressed in the cytoplasm of pachytene spermatocytes. There is no staining in spermatogonia and spermatids. (K) Stage X - XI seminiferous tubules of mouse testis. BOL expression peaks in late pachytene stage spermatocytes. (L-M) Lower magnification view of staining with preimmune and anti-BOL antisera. (L) Preimmune of BOL antisera (magnification:

200X), (M) BOL antisera (magnification: 200X). Spg—spermatogonial cells, Spc—spermatocytes, SSpc—secondary spermatocytes, Spd—spermatids, Pgc—primordial germ cell. DAZ and DAZL antisera are described (22,26).

Figure 3. Specific interaction of BOL and DAZ. (A) The DAZ protein coimmunoprecipitated with BOL when HA beads were incubated with mammalian supernatant from cells expressing both DAZ and BOL fused to an HA tag; shown is DAZ protein detected by Western analysis. DAZ protein was not coimmunoprecipitated with HA beads in the presence of mammalian supernatant from cells expressing DAZ or BOL alone. Right lane shows supernatant from mammalian cells expressing DAZ as a positive control. (B) Deletions of DAZ protein define the minimal region of DAZ required for interaction with BOL. The yeast two-hybrid assay was used to test the ability of eleven deletions of the DAZ:GAL4 fusion protein to interact with BOL. Interactions were assayed via the X-GAL filter assay as above. Assays were done as indicated in Methods. +++, Maximal activity equivalent to that of intact interacting proteins; –, no interaction; +/–, weak interaction; +, strong interaction; ++, nearly wild-type interaction.

Figure 4. DAZ/DAZL evolved from Boule. (A) Distance tree based on the protein sequences of human BOL (BOL.hs), fly Boule (Bol.dm), human DAZL (DAZL), and known homologues from mouse (mDazl), worm (cDazl), zebrafish (zDazl), and frog (Xdazl). The numbers indicate changes per residue. DAZ was added to the tree based on a smaller tree comparing only DAZ and DAZL from vertebrates. (B) Multiple sequence alignment of RNP domains in BOL and DAZL homologues. Distance conserved

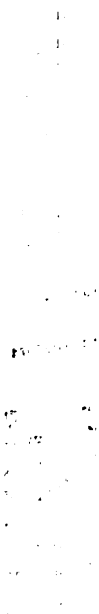
residues are boxed (dark shadow is identical and light shadow is similar). (C) Model of evolutionary history of BOULE/DAZ family. Dark box represents BOL homologues and hatched box represents DAZL homologues. Open box indicates the inferred presence of a BOL homologue that has yet to be identified. The DAZ gene cluster on the Y chromosome is represented by four hatched box linked together to indicate the presence of 4 genes in cluster (4,20). Vertical hatch box indicates the probably period of gene duplication. * represents a duplication event. NWM-new world monkey.

Figure 5. Comparison of functions of members of DAZ/BOL family in different stages of germline development in invertebrates, vertebrates and human lineage with the indicated evolutionary relationship of the members. Expression pattern of each member is indicated by extent of each line. Meiotic expression of *BOL* is seen in fly, mice and humans (primary spermatocytes) and in female oocytes prior to meiosis(7,10). This meiotic function is likely conserved throughout metazoans. DAZL evolved a novel function required for germ stem cell proliferation or differentiation that is unique to vertebrates. DAZL is expressed in multiple compartments from germ stem cells to mature spermatocytes. DAZ arose from DAZL recently in the primate lineage and is expressed in multiple compartments from germ stem cells to meiotic cells. DAZ may enhance gamete quality or quantity production in male primates. OWM—Old world Monkey. Major stages of germcell development were diagramed below. The arrow curve indicates a duplication and horizontal line indicates expression period of each of DAZ/BOL family.

Table 1. Expression and null phenotypes associated with *DAZ/BOL* homologues.

Genes/organisms	Expression			Phenotypes	Ref.
	Testis	ovary	PGC/ Germplasm		
BOULE:					
Drosophila	+	-	-	meiotic arrest	7
Mouse	+	-	-	unknown	*
Human	+	-	-	unknown	*
Worm	-	+	-	meiotic arrest	9
DAZL:					
Zebrafish	+	+	+	unknown	23
Xenopus	+	+	+	early germ cell defect	8,22,10
Mouse	+	+	+	early germ cell defect	6
Human	+	+	+	unknown	17,24,*
DAZ:					
Human	+	-	+	azoospermia & oligospermia	17,24,*

+:protein or mRNA expression observed. -: no expression. *: this paper



References

1. Reijo, R. *et al.* (1995) *Nature Genet.* **10**, 383-393.

2. Reijo, R., Alagappan, R.K., Patrizio, P. & Page, D.C. (1996) *Lancet* **347**, 1290-1293.

3. Shan, Z., Hirschmann, P., Seebacher, T., Edelmann, A., Jauch, A., Morell, J., Urbitsch, P. & Vogt, P.H. (1996) *Hum. Mol. Genet.* **5**, 2005-2011.

4. Saxena, R. *et al.* (1996) *Nature Genet.* **14**, 292-299.

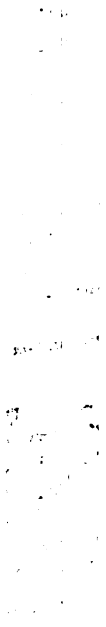
5. Yen, P.H., Chai, N.N. & Salido, E.C. (1996) *Hum. Mol. Genet.* **5**, 2013-2017.

6. Ruggiu, M., Speed, R., Taggart, M., McKay, S.J., Kilanowski, F., Saunders, P., Dorin, J. & Cooke, H. (1997) *Nature* **389**, 73-77.

7. Eberhart, C.G., Maines, J.Z. & Wasserman, S.A. (1996) *Nature* **381**, 783-785.

8. Houston, D.W., Zhang, J., Maines, J.Z., Wasserman, S.A. & King, M.L. (1998) *Devel.* **125**, 171-180.

9. Houston, D.W. & King, M.L. (2000) *Devel.* **127**, 447-456.



10. Karashima, T., Sugimoto, A. & Yamamoto, M. (2000) *Devel.* **127**, 1069-1079.

11. Maegawa, S., Yasuda, K. & Inoue, K. (1999) *Mech. Dev.* **81**, 223-226.

12. Mita, K. & Yamashita, M. (2000) *Mech. Dev.* **94**, 251-255.

13. Maines, J.Z. & Wasserman, S.A. (1999) *Nature Cell Biol.* **1**, 171-174.

14. Ting, C.-T., Tsaur, S.-C., Wu, M.-L. & Su, C.I. (1998) *Science* **282**, 1501-1504.

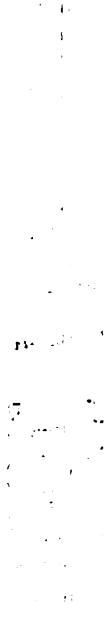
15. Wyckoff, G.J., Wang, W. & Wu, C.-I. (2000) *Nature* **403**, 304-309.

16. Nurminsky, D.I., Nurminskaya, M.V., Aguiar, D.D. & Hardt, D.L. (1998) *Nature* **396**, 572-575.

17. Kent-First, M.G., Kol, S., Muallem, A., Ofir, R., Manor, D., Blazer, S., First, N. & Itskovitz-Eldor, J. (1996) *Mol. Hum. Reprod.* **2**, 943 - 950.

18. Foresta, C., Ferlin, A., Garolla, A., Rossato, M., Barbaux, S. & De Bortoli, A. (1997) *J. Clin. Endocrinol. Metab.* **82**, 1075-1080.

19. Shinka, T. & Nakahori, Y. (1996) *Acta. Paediatr. Jpn.* **38**, 399-404.



20. Saxena, R. *et al.* (2000) *Genomics* **67**, 256-267.

21. Seboun, E. *et al.* (1997) *Genomics* **41**, 227-235.

22. Dorfman, D.M., Genest, D.R. & Pera, R.A.R. (1999) *Hum. Reprod.* **14**, 2531-2536.

23. Chai, N.N., Phillips, A., Fernandez, A. & Yen, P. (1997) *Mol. Hum. Reprod.* **3**, 705-708.

24. Carani, C., Gromoll, J., Brinkworth, M.H., Simoni, M., Weinbauer, G.F. & Nieschlag, E. (1997) *Mol. Hum. Reprod.* **3**, 479-483.

25. Gromoll, J., Weinbauer, G.F., Skaletsky, H., Schlatt, S., Rocchietti-March, M., Page, D.C. & Nieschlag, E. (1999) *Hum. Mol. Genet.* **8**, 2017-2024.

26. Reijo, R.A., Dorfman, D.M., Slee, R., A. A. Renshaw, Loughlin, K.R., Cooke, H. & Page, D.C. (2000) *Biol. Reprod.* **63**, 1490-1496.

27. Cooke, H.J., Lee, M., Kerr, S. & Ruggiu, M. (1996) *Hum. Mol. Genet.* **5**, 513-516.

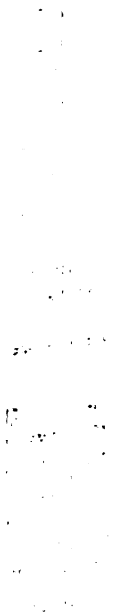
28. Maiwald, R., Luche, R.M. & Epstein, C.J. (1996) *Mamm. Genome* **7**, 628.

29. Menke, D.B., Mutter, G.L. & Page, D.C. (1997) *Am. J. Hum. Genet.* **60**, 237-241.

30. Reijo, R., Seligman, J., Dinulos, M.B., Jaffe, T., Brown, L.G., Disteche, C.M. & Page, D.C. (1996) *Genomics* **35**, 346-352.

31. Cheng, M.H., Maines, J.Z. & Wasserman, S.A. (1998) *Dev. Biol.* **204**, 567-576.

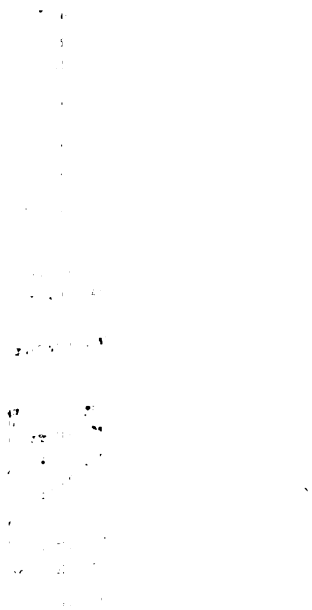
32. Agulnik, A.I., Zharkikh, A., Tong, H.B., Bourgeron, T., K, K.M. & Bishop, C.E. (1998) *Hum. Mol. Genet.* **7**, 1371-1377.

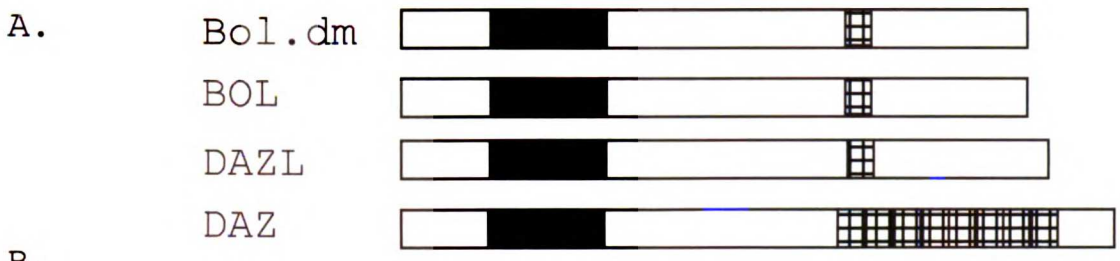


Acknowledgements

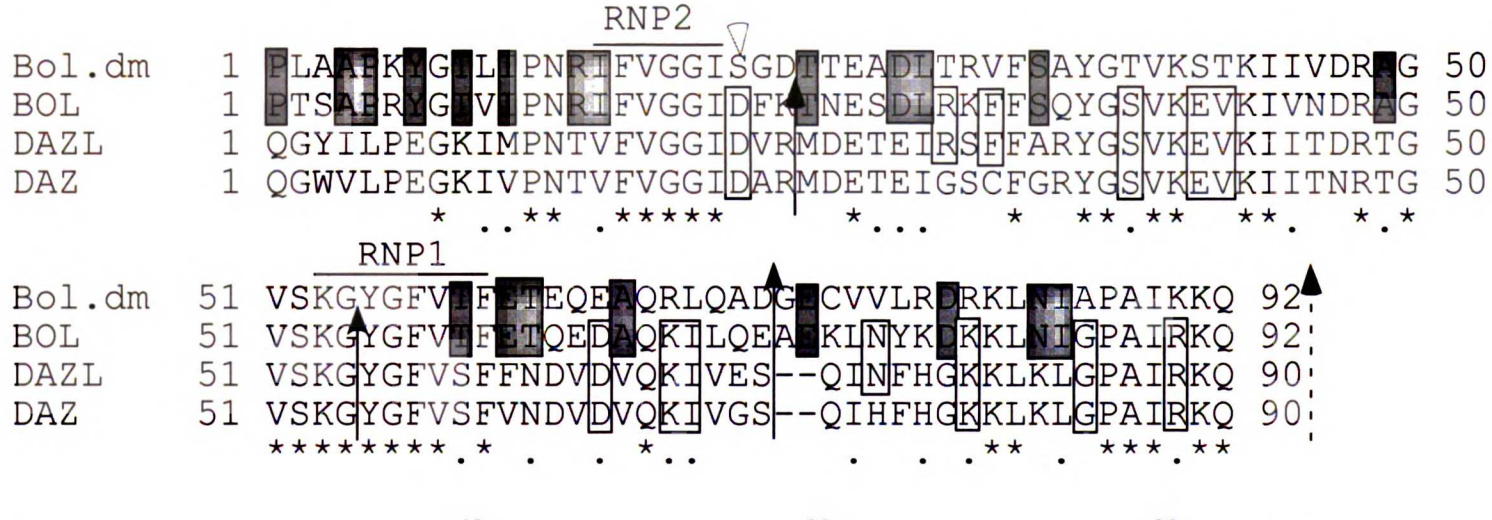
This worked originally appeared in the Proceedings of the National Academy of Sciences Journal in 2001 (vol. 98, p 7414-19). The *BOULE* gene was originally found in our two-hybrid screen and initially characterized by sequence analysis and northern analysis by Frederick Moore. Eugene Xu further characterized the protein sequence of BOL and used phylogenetic analysis to determine the proposed evolutionary model of the *DAZ* gene family. In addition, Eugene Xu determined and characterized the subcellular localization of the BOL protein during spermatogenesis. The interaction between BOL and DAZ proteins was determined and verified by Frederick Moore.

We thank Ira Herskowitz, Holly Ingraham, Synthia Mellon, Robert Taylor, Paul Turek, Richard Weiner, Chung-I Wu, Pauline Yen and members of the Reijo lab for comments on the manuscript, and Jadwiga Jaruzelska for assistance in the two-hybrid screen. This work is supported by a National Research Service Award to E.Y.X., a Ford Foundation fellowship to F.L.M., and grants from the National Institutes of Health, the Searle Foundation and the Sandler Family Foundation to R.A.R.P.




A.
Bol.dm
BOL
DAZL
DAZ

B.


RNP2
RNP1
Bol.dm 1 PLAAHKMGILIPNRIFVGGISGDTTEADLTRVFSAYGTVKSTKIIVDRAG 50
BOL 1 PTSAPRYGTVIPNRIFVGGIDFRTNESDIRKFFSQYGSVKEVKIVNDRAG 50
DAZL 1 QGYILPEGKIMPNTVFVGGIDVRMDETEIRSFFARYGSVKEVKIITDRTG 50
DAZ 1 QGWVLPEGKIVPNTVFVGGIDARMDETEIGSCFGRYGSVKEVKIITNRTG 50
Bol.dm 51 VSKGYGFVTFETQEAAQRLQADGCVVLRDRKLNIAPAIKKQ 92
BOL 51 VSKGYGFVTFETQEDAQKILQEAFKLNYKDKKLNIGPAIRKQ 92
DAZL 51 VSKGYGFVSFFNDVDVQKIVESQINFHGKKLKLGPAIRKQ 90
DAZ 51 VSKGYGFVSFVNDVDVQKIVGSQIHFHGKKLKLGPAIRKQ 90

C.

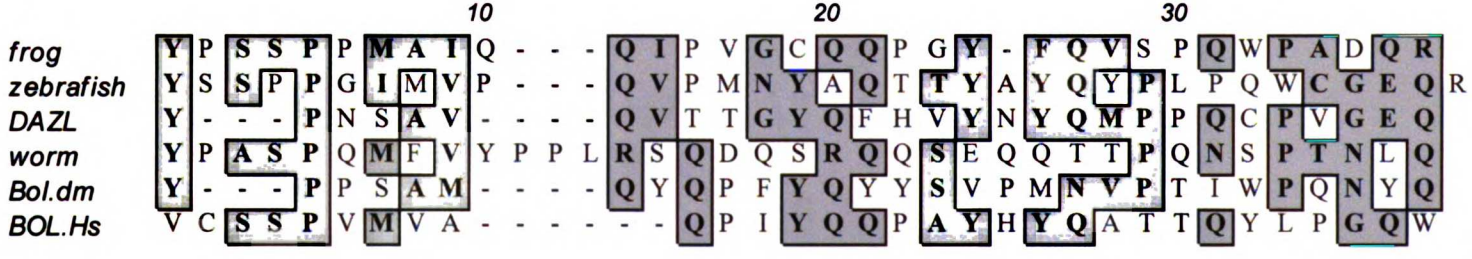

10
20
30
frog
zebrafish
DAZL
worm
Bol.dm
BOL.Hs

Figure 1

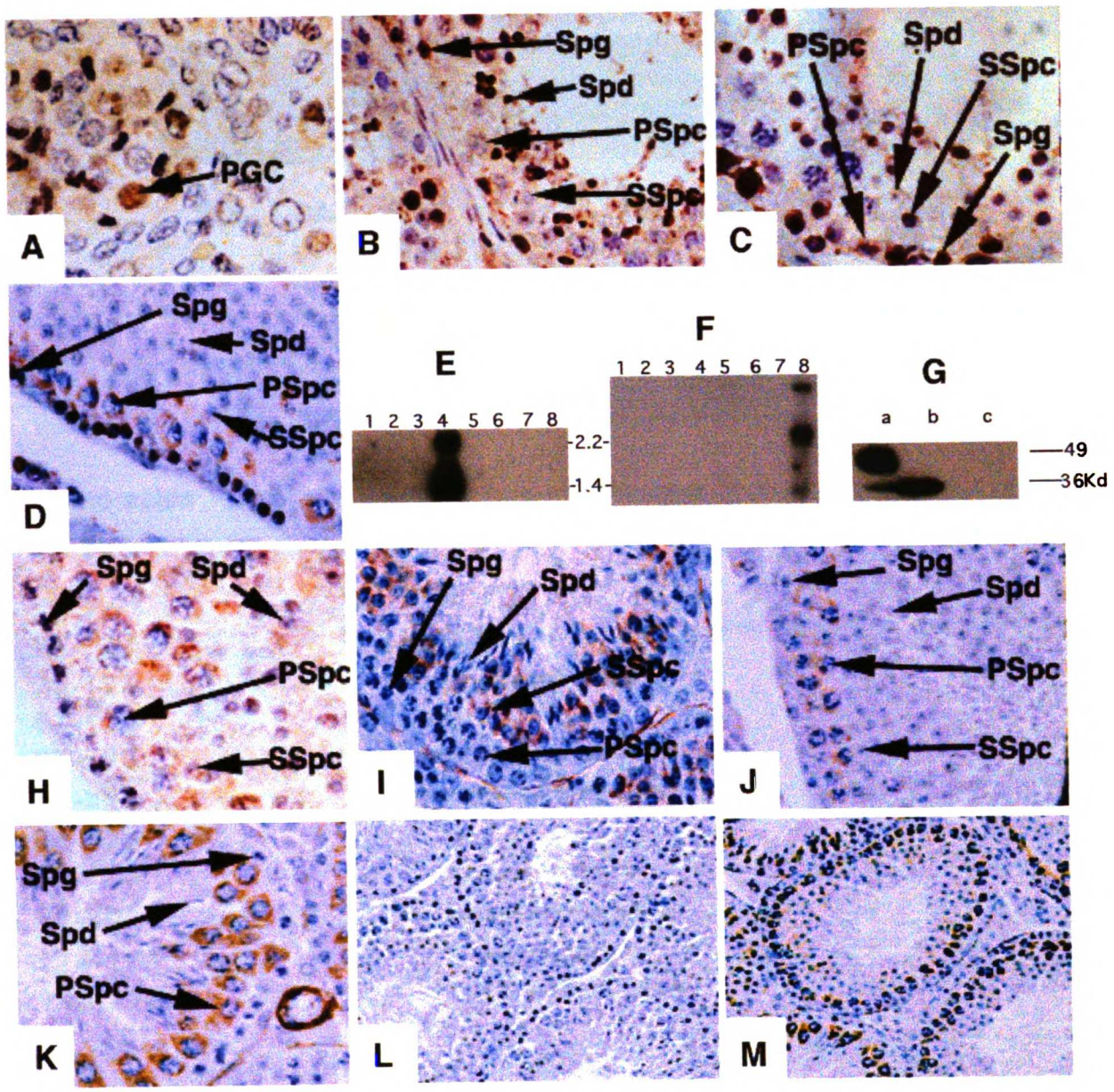
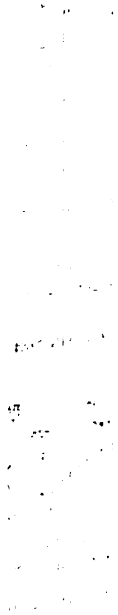

A
PGC
B
Spg
Spd
PSpc
SSpc
C
PSpc
Spd
SSpc
Spg
D
Spg
Spd
PSpc
SSpc
E
1 2 3 4 5 6 7 8
2.2-
1.4-
F
1 2 3 4 5 6 7 8
G
a b c
49
36Kd
H
Spg
Spd
PSpc
SSpc
I
Spg
Spd
SSpc
PSpc
J
Spg
Spd
PSpc
SSpc
K
Spg
Spd
PSpc
L
M

Figure 2



A.

BOL DAZ DAZ DAZ
BOL sup

B.

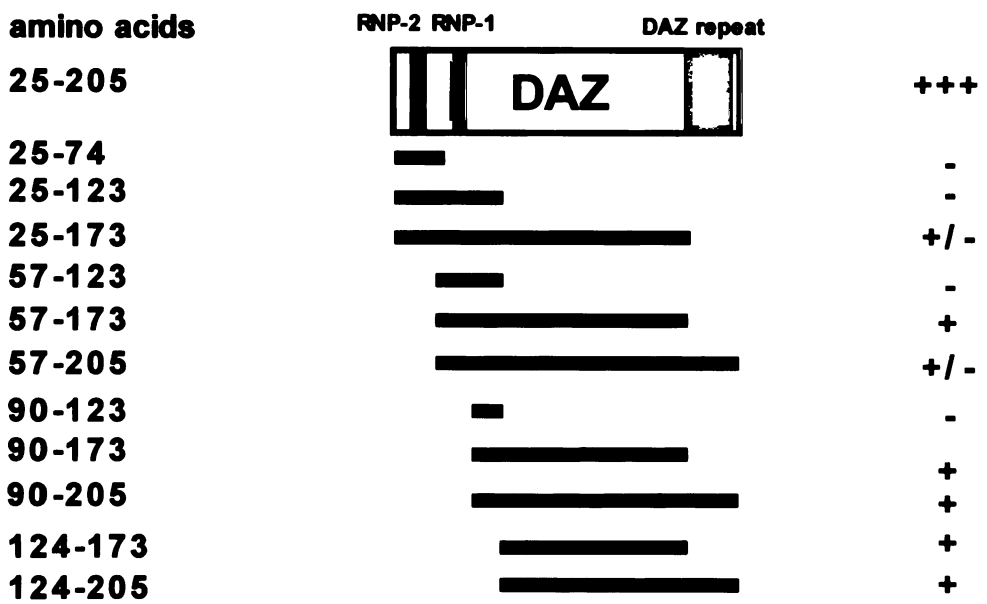
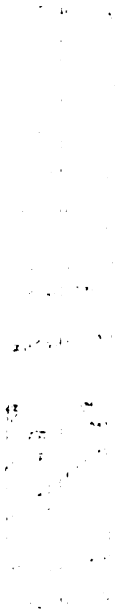


Figure 3



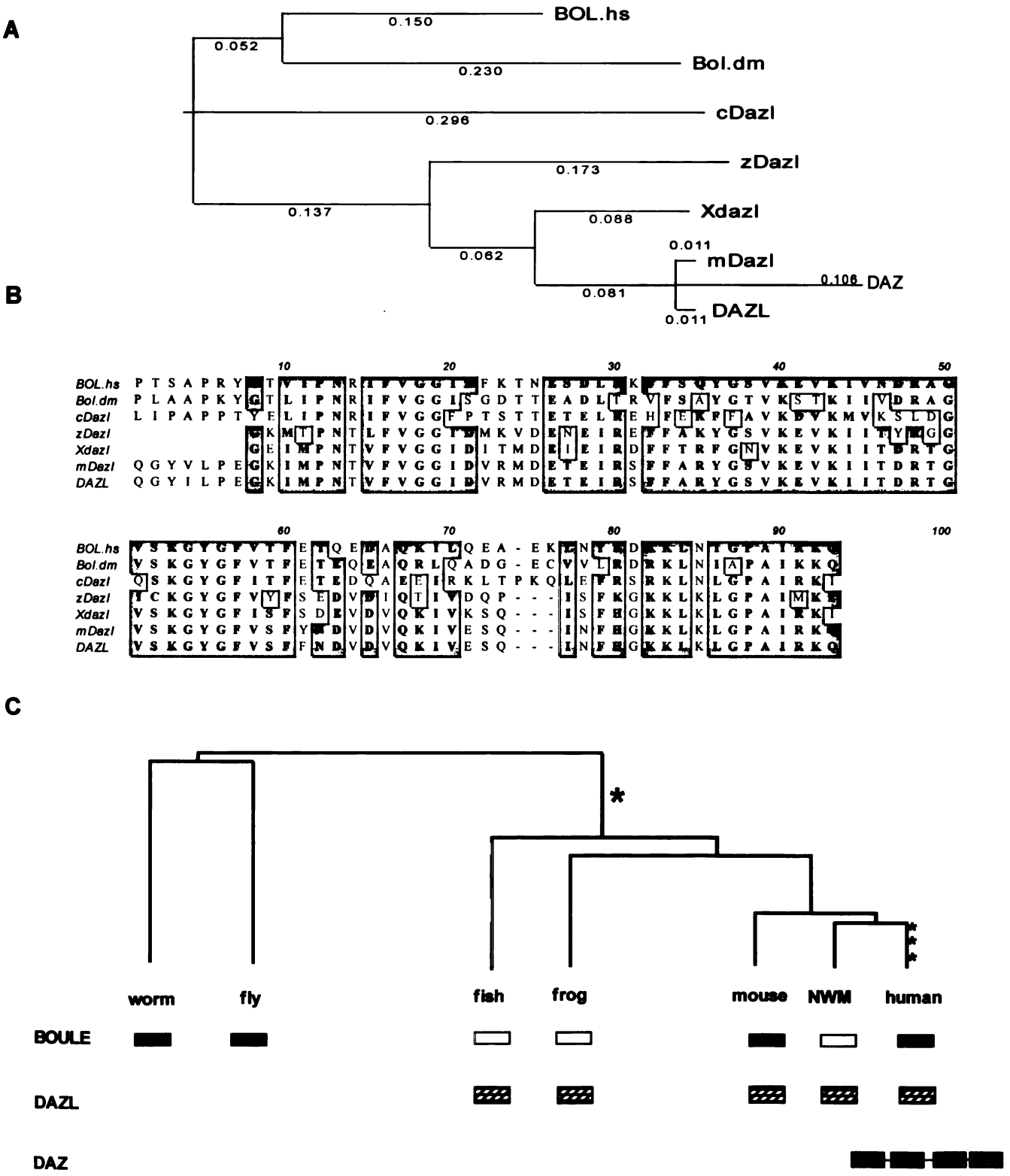

A
BOL.hs
Bol.dm
cDazl
zDazl
Xdazl
mDazl
DAZ
DAZL
0.150
0.052
0.230
0.296
0.173
0.137
0.088
0.011
0.062
0.106
0.081
0.011
B
BOL.hs P T S A P R Y T V T P N R I F V G G I S F K T N E S D L R K
Bol.dm P L A A P K Y G T L I P N R I F V G G I S G D T T E A D L T R V F S A Y G T V K S T K I I V D R A G
cDazl L I P A P P T Y E L I P N R I F V G G F P T S T T E T E L R E H F E K F F A V K D V K M V K S L D G
zDazl G K M T P N T L F V G G I D M K V D E N E I R E F F A K Y G S V K E V K I I T Y E G G
Xdazl G E I M P N T V F V G G I D I T M D E I E I R D F F T R F G N V K E V K I I T D R T G
mDazl Q G Y V L P E G K I M P N T V F V G G I D V R M D E T E I R S F F A R Y G S V K E V K I I T D R T G
DAZL Q G Y I L P E G K I M P N T V F V G G I D V R M D E T E I R S F F A R Y G S V K E V K I I T D R T G
10 20 30 40 50
BOL.hs V S K G Y G F V T F E T Q E E A Q K I L Q E A - E K L N Y S D L K L N I G P A I K K Q
Bol.dm V S K G Y G F V T F E T E Q E A Q R L Q A D G - E C V V L R D R K L N I A P A I K K Q
cDazl Q S K G Y G F I T F E T E D Q A E E R K L T P K Q L E F R S R K L N L G P A I R K T
zDazl I C K G Y G F V Y F S E D V D I Q T I V D Q P - - - I S F K G K K L K L G P A I M K E
Xdazl V S K G Y G F I S F S D E V D V Q K I V K S Q - - - I S F H G K K L K L G P A I R K T
mDazl V S K G Y G F V S F Y N D V D V Q K I V E S Q - - - I N F H G K K L K L G P A I R K E
DAZL V S K G Y G F V S F F N D V D V Q K I V E S Q - - - I N F H G K K L K L G P A I R K Q
60 70 80 90 100
C
*
worm
fly
fish
frog
mouse
NWM
human
BOULE
DAZL
DAZ

Figure 4

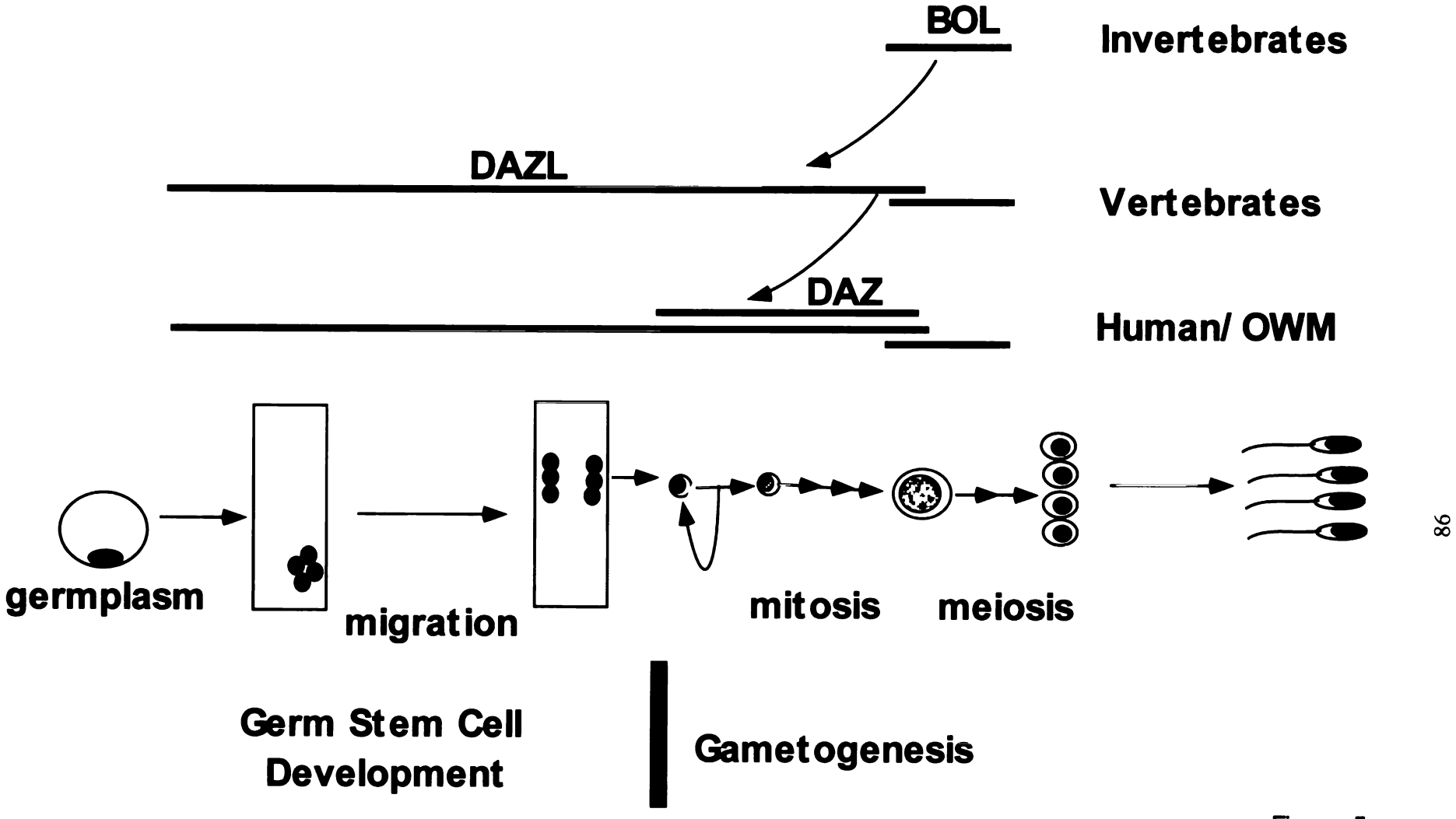

BOL
Invertebrates
DAZL
Vertebrates
DAZ
Human/ OWM
germplasm
migration
mitosis
meiosis
Germ Stem Cell Development
Gametogenesis

Figure 5

Chapter 4

Zinc-finger Proteins, DZIP (DAZ Interacting Protein) 1 & 2, Associate with DAZ (Deleted in AZoospermia) and are Implicated in Embryonic and Germ Cell-Specific RNA-binding Complexes

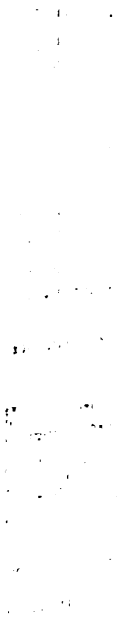
Zinc-finger Proteins, DZIP (DAZ Interacting Protein) 1 & 2, Associate with DAZ (Deleted in AZoospermia) and are Implicated in Embryonic and Germ Cell-Specific RNA-binding Complexes

Frederick L. Moore*, Jadwiga Jaruzelska*†, David M. Dorfman¥, and Renee A. Reijo-Pera*

*Department of Obstetrics, Gynecology and Reproductive Sciences; Department of Physiology; Department of Urology; and Programs in Human Genetics and Cancer Genetics; University of California at San Francisco, San Francisco, CA 94143-0546, USA

†Institute of Human Genetics, Polish Academy of Sciences, Poznan, Poland

¥Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, 75 Francis Street, Boston, MA 02115, USA



Abstract

During embyrogenesis, both the translation of maternal transcripts and the activation of zygotic genes are required to allocate or set aside cells destined for the germ cell lineage[1]. Evidence in humans and model organisms suggests that the Y-linked *DAZ* (*Deleted in AZoospermia*) genes and autosomal homologs *DAZL* (*DAZ-Like*) may play a role early in germ cell development to maintain initial germ cell population[2-8]. The demonstration that DAZ and DAZL proteins associate with PUM2[9], a human homolog of the *Pumilio* gene in *Drosophila* which plays a role in germline stem cell maintenance and functions during embryogenesis and primordial germ cell migration as a translational repressor[9, 10], has begun the process of characterizing the molecular function of DAZ and DAZL proteins. Here we report the identification and characterization of the *DZIP (DAZ Interacting Protein) 1 & 2 genes*, which encode closely related proteins that contain a C2H2 zinc-finger domain and associate with DAZ protein. Both the *DZIP1* and *DZIP2* genes are expressed predominantly in human embryonic stem cells and in the germ cells of the fetal ovary and testis and adult testis; moreover, DZIP1 colocalizes with DAZ and DAZL proteins in fetal germ cells and adult testis at the cellular and subcellular levels. Characterization of the regions of interaction between DAZ and DZIP2 proteins suggests that DAZ and its associated cofactors may form an RNA-binding protein complex.

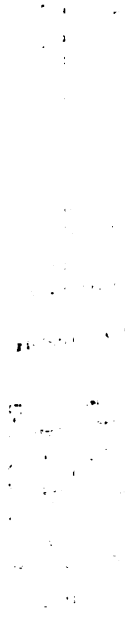
Introduction

The requirement for maternal transcripts early in embryogenesis to designate the germ cell lineage is evident in model organisms by mutations and deletions of genes that affect the formation of germ plasm and subsequent primordial germ cells. Shortly after embryogenesis begins, transcription and translation of zygotic genes necessary for the migration, proliferation, and maintenance of cells in the germ cell lineage are also required. In humans, evidence suggests that the *DAZ* (*Deleted in AZoospermia*) genes are strong candidates for factors that play a role early in germ cell development to maintain the initial germ cell population. Deletions of the Y-linked *DAZ* gene cluster result in the complete loss of germ cells and/or a severe reduction in the number of germ cells in the testis[2, 11]. Three members of the *DAZ* gene family exist in humans, and the homologs of these genes in model organisms suggest an ancient conserved function in germ cell development. The *BOL* gene, a homolog of a meiotic regulator in *Drosophila*, is expressed at the onset of meiosis in adult males[12]. Disruption of the *boule* gene in *Drosophila* causes maturing germ cells in male flies, to arrest at the pachytene stage of meiosis, and causes sterility[13-15]. Secondly, the *DAZL* (*DAZ-Like*) gene, located on chromosome 3, is expressed in the germ cells of male and females[16-18]. Evidence suggests that the *DAZL* gene arose from a duplication of the *BOL* gene[12]. The expression profile of *DAZL* early in fetal gonocytes through spermatogenesis and oogenesis suggests it acquired additional functions[5, 6]. Lastly, the *DAZ* gene, which is located in four copies on the Y chromosome, is expressed early in gonocytes through spermatogenesis[2, 6, 19]. Evidence suggests that the *DAZ* gene arose from a duplication

of *DAZL* and is present in great apes and old world monkeys[20-22] and not present in new world monkeys and other mammals.

The function of *DAZ* and *DAZL* early in germ cell development is supported by experiments with *DAZ* homologs from model organisms. Loss of function of *Dazl* in the mouse results in sterility in male and female mice due to depletion in the germ cell population[3]. A partial rescue experiment with a *DAZ* transgene in the *Dazl* knockout mouse supports a conserved function of DAZ and DAZL[23]. The *DAZ* homolog in *Xenopus*, *Xdazl*, has been shown to function early in germ cell development; Inhibition of *Xdazl* leads to loss of their primordial germ cells[4, 24].

The molecular function of DAZ and DAZL proteins has been alluded to by their association with PUM2, which is a human homolog of the *Pumilio* gene in *Drosophila* that functions during embyrogenesis and germ cell development, and is required to maintain the germline stem cells[9, 10, 25-29]. Pumilio in *Drosophila* is an RNA-binding protein that associates with a combination of other cofactors at the 3' UTR region of specific transcripts during embryogenesis and primordial germ cell migration and represses translation of transcripts required for development[25]. The DAZ protein has been shown to interact with PUM2 in the same region as Pumilio cofactors in *Drosophila* and form a ternary RNA-protein complex[9]. We report the identification and characterization of two novel zinc-finger proteins, called DZIP1 (DAZ Interacting Protein) & 2, which interact with DAZ protein and are expressed early in embryonic stem cells and maintained in fetal gonocytes and male germ cells.



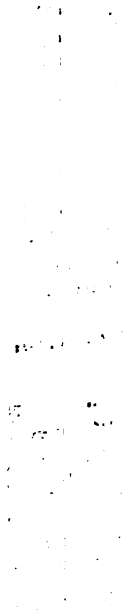
Material & Methods

Identification of DZIP1 and DZIP2

The yeast two-hybrid system was used to screen for proteins that interact with DAZ proteins expressed as a fusion with the GAL4 DNA-binding domain (Clontech, Inc., Palo Alto, CA). This construct was derived from a cDNA that contains the N-terminal RNA-binding domain and a single DAZ repeat (pRR102). A testis library of random cDNAs fused to the activation domain of GAL4 was transformed into yeast and screened according to the manufacturer's instructions. *DZIP1* and *DZIP2* cDNAs were isolated from our screen six and two times, respectively. The DZIP1 and DZIP2 proteins interacted specifically with DAZ protein as demonstrated by the observations that they did not interact with the DNA-binding domain alone or with a construct with the DNA-binding domain fused to laminin. In addition, the DZIP1 and DZIP2 proteins did not activate transcription on their own. Sequencing of *DZIP1* and *DZIP2* clones was done at the Genome Core Facility of the Program in Human Genetics (UCSF).

Expression of *DZIP1* and *DZIP2* transcripts

Northern hybridization was as described[2]. Primers used to produce probes for Northern analysis were: DZIP1F, TGGAAGATCCTTTTCCCAGA; DZIP1R, AGTGTCCTCGATTTCGCTTC; DZIP2F, TTTGGTCATCTTTGAGATTCTGAC; DZIP2R, CAGATGTACACATGACTGAATGGA. For RT-PCR analysis, fetal tissues and biopsy tissue from adult men were homogenized in Trizol reagent according to



manufacturer's instructions (Gibco-BRL, Inc.; Bethesda, MD). Then 1 μg of total RNA was reverse-transcribed using oligo-dT primers (Roche, Inc.). Following reverse transcription, 100 ng of cDNA was subjected to PCR with DZIP1 or DZIP2 primers. Sequences were: DZIP1F, GCAAGCTGGAAGACGAGAAG; DZIP1R, ACTGGAGCTGTCCATTTGCT; DZIP2F, GCAAGCTGGAAGACGAGAAG; DZIP2R, CAGATGTACACATGACTGAATGGA.

Coimmunoprecipitation of DZIP1 and DZIP2 with DAZ

For coimmunoprecipitation, yeast cells expressing DAZ alone, DZIP1 or DZIP2 alone and both DAZ and DZIP1 or DZIP2 were pelleted, lysed and resuspended in lysis buffer (50 mM Tris-HCl, pH 7.5, 1 mM EDTA, 150 mM NaCl, 10% glycerol, 1% NP-40) with a cocktail of protease inhibitors (Clontech, Inc.). The *DZIP1* and DZIP*2* cDNAs, which are fused to the coding sequence for the GAL4 activation domain that contains a hemagglutinin (HA) tag, were used to immunoprecipitate DZIP1 and DZIP2 from yeast supernatants. HA antibodies covalently linked to inert beads (BABCO, Inc.) were incubated with yeast supernatants, washed with lysis buffer 3 times, resuspended in 1 mM glycine (pH 2.8), and incubated at 30°C to elute bound proteins. Proteins were separated on a sodium dodecyl sulfate-polyacrylamide gel. The presence of DAZ proteins was assessed by Western analysis with DAZ-specific antibodies [5, 6].

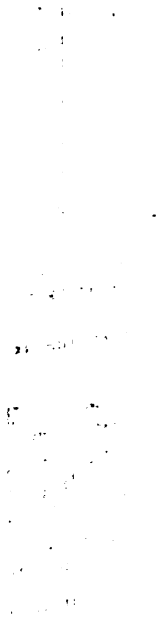
Deletion analysis

Deletions of the open reading frames of DZIP2 were constructed using primers designed to clone fragments of *DZIP 2* cDNAs in frame with the GAL4 activation domain within the pACT2 vector. The constructs were transformed into Y190 yeast strain

along with the DAZ:GAL4 DNA-binding fusion construct and screened for protein-protein interactions by X-Gal filter assays (Clontech, Inc.). Likewise, for DAZ deletions, primers were used to clone fragments of the *DAZ* cDNA in frame with the GAL4 DNA-binding domain within the pAS2 vector (Clontech Inc.). The constructs were transformed into yeast along with DZIP2:GAL4 activation domain fusion construct and screened for interaction.

Expression analysis

For Western blotting and immunocytochemistry, polyclonal antibodies to DZIP1 were raised in rabbits by injection of the synthetic oligopeptide, SSHILEPIEELSEEEKGREN, coupled to KLH (Research Genetics, Inc., Huntsville, AL). This peptide is only found in DZIP1; it is not present in DZIP2. DZIP1 antisera were collected after 10 weeks. Western blotting and immunocytochemistry were done as described [6]. Human testis sections were obtained from a 20-21 week fetus and from a 38-year-old man who presented with a seminoma; the sections were not in contact with the seminoma. Human female tissue sections were from fetal ovary, 19 weeks gestation. Mouse testis sections were obtained from 60-day adult mice. DZIP1 antisera and preimmune control antisera were both used at dilutions of 1:100 and 1:200 for immunocytochemistry and 1:1000 for Western blots. Localization was determined by enlarging the images and examining the position, size, and morphology of cells that were determined immunopositive as previously described[12].



Electron microscopy

The testis samples were washed in PBS once before fixed (2% para-formaldehyde / 0.1% glutaraldehyde / PBS) for 2-4 hours at room temp., stored in PBS at 4°C overnight, and processed for LR White embedding and immunolabelling. Embedding was done at -20°C and cold curried with the accelerator (LR White)[30]. Silver to gold sections were cut with Reichert-Jung Ultracut E ultra-microtome, mounted on 300 mesh naked nickel grids, and air dried overnight. Grids were pretreated for 10 minutes with saturated aqueous m-sodium periodase (Sigma, St. Luis, MI.) followed by 10 minutes 0.1N HCl, and blocked with 2% fish gelatin and BSA in PBS. Grids were incubated with affinity-purified antibodies for 1-2 hours, and incubated for 30 minutes with goat anti-rabbit IgG conjugated with 10 nm gold (Amersham Life Science, Arlington Heights, Ill.). The grids were fixed briefly with glutaraldehyde and osmium tetraoxide, stained with uranyl acetate and examined with a Philips Tec Nai 10 electron microscope.

Results

Identification of DZIP1 and DZIP2

The DAZ protein was used as bait in a two-hybrid screen to identify transcripts in the testis, which express proteins that interact with DAZ. Seven candidate genes were isolated from our screen[9]. The genes obtained most frequently were related and encoded novel zinc-finger genes that we called the *DZIP* (*DAZ Interacting Proteins*) *1 & 2*. Transcripts of *DZIP1* and *DZIP2* were isolated from the two-hybrid screen six times and two times, respectively. Comparison of *DZIP1* and *DZIP2* cDNAs revealed that the majority of their nucleotides are identical with major differences being two insertions of ~140 nucleotides and a unique 3' UTR region of the *DZIP2* transcript (Fig. 1a). Theoretical translation of the *DZIP2* cDNA indicated that the first insert results in a 46 amino acid insertion at the N-terminus of DZIP2 and the second insert results in a truncated protein shortly after the zinc-finger domain, when compared to theoretical translation of the *DZIP1* transcript (Fig. 1a, b). To determine the chromosomal location of *DZIP1* and*DZIP2*, we searched the GenBank database using the unique 3' UTR region of both transcripts, which revealed that *DZIP1* and *DZIP2* are located on chromosomes 13q31 and 3p13, respectively.

Analysis of *DZIP1* and *DZIP2* Expression

We next sought to characterize the expression of *DZIP1* and *DZIP2* in various adult and fetal tissues. Northern analysis using a probe that was specific to *DZIP1* detected three transcripts of ~ 1.35, 3.9, and 4.4 kilobases in the testis (Fig. 2a).

Interestingly, the *DZIP1* gene is predominantly expressed in the testis and at lower levels in skeletal muscle, ovary, and heart. Northern analysis using a probe that is specific to *DZIP2* detected a transcript of ~3.9 kilobases in the testis (Fig. 2b). Northern blots for *DZIP2* were exposed 3 times longer compared to *DZIP1*, yet signals were not detected in other tissues (Fig. 2b), including a Northern blot containing mRNAs from heart, brain, placenta, lung, liver, skeletal muscle, kidney, and pancreas tissues (data not shown). We next sought to determine if *DZIP1* and *DZIP2* genes are expressed early in human development by RT-PCR on cDNA from human embryonic stem cells (ES), fetal tissue, and adult tissue. Using primers specific to *DZIP1*, transcripts are detected in human ES cells, fetal ovary and testis, and adult testis (Fig 2c). Likewise, RT-PCR was used with *DZIP2*-specific primers and detected mRNA expression in human ES cell, fetal brain, fetal ovary, fetal testis, and adult testis (Fig 2c).

Analysis of DZIP1 and DZIP2 interaction with DAZ protein

If the interaction of DAZ protein with DZIP1 and DZIP2 proteins is legitimate, we expect that they should form a complex that can be biochemically isolated by co-immunoprecipitation. To test this, yeast supernatant was derived from yeast expressing DAZ alone, DZIP1 or 2 alone, and both DAZ and DZIP1 or 2. Then, we used an HA-tagged DZIP1 and DZIP2 fusion protein to isolate DZIP proteins from yeast supernatant and probed by Western analysis for the presence of DAZ protein. Our results show that the DAZ protein is only isolated from the yeast supernatant containing both DAZ and DZIP proteins (Fig. 3a, lane 3, 8). These data show that the DAZ protein can form a stable complex with DZIP1 and DZIP2 proteins.

To define the minimal regions necessary for DAZ and DZIP interactions, we used deletion analysis with various constructs in the yeast two-hybrid system. Analysis of cDNA deletions that produce truncated DAZ protein indicated that the minimal region of DAZ necessary for interaction with DZIP2 is between amino acid residues 90 and 123. Interestingly, when deletions that remove the DAZ repeat, beginning at amino acid 25 (Fig. 3b) and amino acid 57 (data not shown), totally abolish interaction with DZIP2. Also of interest, truncated DAZ protein that begin at amino acid 90, which is just outside the RNA-binding domain, interact strongly with DZIP2. In addition, the DAZ protein that begins at amino acid 90 and also contains a DAZ repeat (amino acid 90-205) interacts with DZIP2 stronger than wild type (Fig. 3b). Taken together, these data suggest that DZIP2 may interact with two regions of DAZ, the region outside of the RNA-binding domain between amino acid 90-123 and with the DAZ repeat. Truncated forms of DAZ protein that disrupt folding of the RNA-binding domain and do not contain the DAZ repeat have reduced or no interaction with DZIP2 protein. We next created cDNA deletions that produce truncated DZIP2 protein to analyze the minimal region necessary to interact with DAZ. Deletion analysis of DZIP2 indicated that DAZ interacts in the zinc-finger region of DZIP2 (Fig. 3c).

Localization of DZIP1 protein in fetal and adult germ cells

The DAZ and DAZL proteins localized to distinct compartments in adult testis and ovaries[5, 6, 31, 32]. We next sought to determine if DZIP1 protein colocalizes with DAZ and DAZL in germ cells. Antisera was generated against a DZIP1-specific peptide

in rabbits and used to probe a Western blot of extracts from human testis (Fig. 4a, lane 1,4), mouse testis (lane 2,5), and mouse embryonic stem cells (lane 3,6). Anti-DZIP serum detected protein bands in human testis of ~ 89 and 123 kDs (lane 1), in mouse testis of ~80 and 120 kDs (lane 2), and in mouse embryonic stem cells of 92 and 123 kDs (lane 3). Translation of the *DZIP1* cDNA isolated from our two-hybrid screen produced a theoretical protein of 89 kD (accession # AF272347). An EST deposited in Genbank indicates an isoform of DZIP1 with a theoretical size of 96 kD (accession # NM_014934). Results from our Northern and Western analysis suggest that several isoforms of DZIP1 protein may exist. Although we also attempted to generate antisera specific to the DZIP2 protein by using a peptide in the insert region at the N-terminus of the DZIP2 protein, the DZIP2 antiserum were not specific. To determine if expression of *DZIP2* is specific to germ cells or somatic cells present in the testis, we thus resorted to RT-PCR analysis of cDNA from the testis of normal men (with germ cells present) and men with Sertoli cell only syndrome (no germ cells present). Results indicate that *DZIP2* message is only expressed in germ cells and is not present in the testis of patients who lack germ cells (Fig. 4b).

The DAZ and DAZL proteins are present in male germ cells at several stages in their development, both prenatally and during spermatogenesis[6]. Thus, we sought to determine if the DZIP1 protein is localized at the same time and place. Fetal testis sections were stained with DZIP1 antibody serum and indicate that DZIP1 localizes to the nucleus and cytoplasm in primordial germ cells (PGCs), which is identical to DAZ and DAZL proteins[6] (Fig. 4d). In the adult testis, DAZ and DAZL proteins have been

shown to localize to the nucleus and cytoplasm in spermatogonia cells and relocate at the onset of meiosis exclusively to the cytoplasm in spermatocytes[6, 32]. Staining of DZIP1 in human testis indicated identical localization of DZIP1 to the nucleus and cytoplasm of spermatogonia and cytoplasmic staining in spermatocytes (Fig 4f). Our RT-PCR results indicated that the *DZIP1* transcript is also present in fetal ovaries and staining in this tissue verifies that DZIP1 protein is also present in PGCs of the fetal ovaries (Fig. 4h), similar to DAZL[5].

The localization of DZIP1 protein in mouse testis mirrors that of humans: immunostaining of mouse testis sections for DZIP1 protein showed identical staining in spermatogonia and spermatocytes (Fig. 5a, b). Interestingly, DZIP1 protein localizes to a distinct U-shaped pattern in spermtocytes (Fig. 5a, b). We next sought to determine the localization of DZIP1 protein at the subcellular level in late stage spermatocytes and round spermatids. Ultrastructural analysis in mouse testis confirms DZIP1 cytoplasmic localization in spermatocytes (data not shown) and also detects a high concentration of DZIP1 in the pro-acrosome of early spermatids (Fig. 5d). We also detect a high concentration of DZIP1 in the acrosomal region of late round spermatids (Fig. 5e). In addition, DZIP1 staining persists in the cytoplasm and acrosomal head region of elongating spermatozoa (Fig. 5f). Ultrastructure analysis of Dazl was inconclusive. However, immmunofluorescence and histochemistry suggests that it too is not evenly distributed throughout the cytoplasm[5, 6, 9].

Discussion

DZIP1 and DZIP2 Sequence and Structure

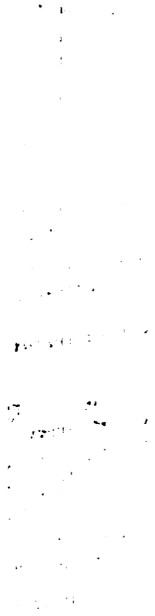
The nucleotide sequences of the *DZIP1* and *DZIP2* cDNA transcripts are strikingly similar even though the *DZIP1* and *DZIP2* genes are located on different chromosomes. The sizes of the *DZIP1* and *DZIP2* transcripts isolated from our two-hybrid screen are 2.4 and 2.56 kilobases, respectively. The first 1817 nucleotides of the *DZIP2* transcript are identical to the *DZIP1* transcript, except for two insertions in the *DZIP2* transcript that results in an additional 46 amino acid in the N-terminus of DZIP2 and a truncation of DZIP2 protein shortly after the C2H2 zinc-finger domain. The 3' UTR region of *DZIP2* after the second insert is also identical to *DZIP1* up until nucleotide 1817; thereafter, the 3' UTR region of *DZIP2* is unique. The similar nucleotide sequences suggest that the *DZIP1* and *DZIP2* genes arose from a duplication event. Sequence analysis suggests that the genes are novel; similar sequences are not found in *Drosophila* or *C. elegans*, but thus far the *Dzip1* gene has been found in mice. Thus, it is likely that the duplication of the *DZIP* genes occurred during vertebrate evolution. This conclusion is strengthened by examination of the expression profile of *DZIP1* and *DZIP2* transcripts; expression is nearly identical. Both *DZIP1* and *DZIP2* transcripts are detectable in ES cells, fetal testis and ovaries, and adult testis. Additional analysis of the genomic sequence of *DZIP1* and *DZIP2* and comparison with DZIP1 and DZIP2 homologs in other organisms will address this point. Also of interest, when we analyzed the composition of the 46 amino acids unique to the DZIP2 protein, it indicated 14 prolines and 7 glycines. Proteins with sequence containing proline-rich domains have

been indicative of roles in signal transduction. Deletion analysis indicated that the 46 amino acid region is not necessary for DAZ interaction with DZIP2, suggesting that this region may have a novel function unique to DZIP2.

The function of known DAZ-associated cofactors

The identification of previously characterized genes that encode proteins that interact with DAZ and DAZL has allowed suggestions for possible functions for members of the *DAZ* gene family. For example, PUM2, a human homolog of the well characterized *Pumilio* gene in *Drosophila*, which functions to repress translation of transcripts required for primordial germ cell migration and also regulates translation of transcripts required for germline stem cell maintenance[25], forms a stable complex with DAZ and DAZL proteins[9]. The Pumilio protein in *Drosophila* interacts with a combination of different cofactors, one of which is a zinc-finger protein, as it binds to specific sites in the 3' UTR region of transcripts that it regulates[33, 34]. In humans, DAZ and DAZL proteins interact with PUM2 in the same region and bind the same RNA target as Pumilio and its cofactors in *Drosophila*[9].

In addition, BOL, which has been characterized in *Drosophila* as a testis-specific regulator of meiosis entry and germline differentiation[15], has been shown to interact with the DAZ protein (Moore and Reijo, unpublished data). Evidence in *Drosophila* indicates that *boule* interacts genetically with *twine*, a meiotic cell division cycle protein Cdc-25 phosphatase, and is required for the translational activation of the *twine* transcript at the onset of meiosis[15]. Disruption of the *boule* gene in *Drosophila* results in arrest



of germ cells and sterility in male flies[14]. The molecular function of BOL is suggested to be conserved from humans to flies based on transgenic experiments in flies, where the human BOL can partially rescue the meiotic arrest phenotype in *boule* flies (Xu, unpublished data). The molecular role of DZIP1 and DZIP2 proteins remains to be elucidated, but their association with the DAZ protein and their capacity to bind RNA based on their C2H2 domain, does not exclude a possible role in translational regulation.

DAZ and Associated Cofactors Interacting Regions

Analysis of the various regions of the DAZ protein that interact with associated cofactors suggests that DAZ may function as a part of a protein complex. Previously, the DAZ and DAZL proteins have been shown to hetrodimerize with each other and also homodimerize with itself[35, 36]. The minimal region necessary for Dazl-Dazl interaction in the mouse is between amino acids 80-132 just C-terminus to the RRM domain[35]. Yet the multiple DAZ repeats located at the C- terminus of the DAZ protein have been shown to interact in DAZ-DAZ dimerization[36]. These results taken together with the isolation of DAZL from our two-hybrid screen using DAZ protein with one DAZ repeat suggest that the minimal region necessary for DAZ and DAZL homo- and heterodimerization is the region C-terminal to the RRM domain. In addition, the DAZ repeats also play a role in protein-protein interaction. Deletion analysis using an intact DZIP2 protein and a series of truncated DAZ proteins indicated that the minimal region necessary for DAZ interaction with DZIP2 is the region C-terminal to the RRM domain between amino acids 90-123. Notably, the truncated DAZ protein that contained amino acids 90 and 205, and one DAZ repeat interacted with DZIP2 similar to intact DAZ.

These results indicate that DZIP2, and likely DZIP1, interact with DAZ protein in the same region that DAZ and DAZL proteins form homo- and heterodimers. Also of interest, the minimal region of DAZ required for interaction with PUM2 protein is the linker region between amino acids 124-173 [9], which excludes the DAZ repeats and DZIP2 binding region. Moreover, two DAZ and DAZL interacting cofactors, DAZAP (DAZ Associated Proteins) 1 and 2, have been shown to utilize the DAZ repeats for interaction[36]. Based on the unique regions of DAZ that interact with DZIP2, PUM2, DAZAP1 and DAZAP2 proteins, it is tempting to speculate that DAZ and DAZL proteins are capable of forming protein clusters with a combination of different cofactors on RNA targets that they regulate.

Localization of DAZ and DAZL in Relations to their Cofactors

The DAZ and DAZL proteins have been previously detected in the primordial germ cells of fetal ovary and/or testis[5, 6]; moreover, evidence from deletions of DAZ and its homologues in human and mouse suggests a function early in germ cell development for the maintenance of germ cells[2, 3]. Our research indicates that both *DZIP1* and *DZIP2* transcripts are present early in ES cells and that DZIP1 expression is limited to fetal testis and ovaries and adult testis, and similarly that DZIP2 expression is limited to fetal tissue (brain, testis, and ovary) and adult testis. With the addition of localization data for DZIP1, which is identical to that of DAZ and DAZL proteins in fetal germ cells and adult testis, both DZIP1 and DZIP2 proteins can likely function early in germ cell development in conjunction with DAZ and DAZL proteins to maintain the population of germ cells. Interestingly, the expression profile of DZIP1 and DZIP2 is

similar to the DAZ-associated cofactor PUM2. The *PUM2* gene is expressed in ES cells, and fetal and adult germ cells and shares identical staining with DAZ and/or DAZL proteins in those tissues[9]. Thus, DAZ and DAZL may associate with a combination of cofactors, some of which are germ cell-specific and others that are expressed ubiquitously, and regulate translation of transcripts early in germ cells development; thereafter, DAZ and DAZL proteins may interact with additional cofactors at meiosis to allow the maturation of spermatocytes.

Conclusion

Additional experiments that will address the loss of function of *DZIP1* and/or *DZIP2* genes in the mouse will help to further elucidate the function of the *DZIP* gene family early in germ cell development and through spermatogenesis. Molecular experiments that address whether DZIP1 and DZIP2 associate with DAZ and additional cofactors and the identification of transcripts that they bind and possibly regulate will be the next step in unraveling the germ cell development pathway in humans at the molecular level. The identification of DAZ-associated proteins, like DZIP1 and DZIP2, where their expression suggests a function unique to ES cells and germ cells, allows science community to have additional tools to mark germ cells and develop innovative ways to dissect the germ cell pathway.

Figure Legends

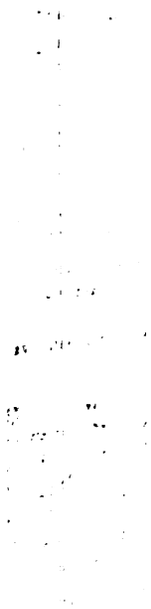
Figure 1 DZIP1 & 2 sequence analysis, a, Schematic of nucleotide alignment of *DZIP1* and *DZIP2* cDNA sequence. Horizontal black lines represent identical nucleotides and hatched boxed area represents unique nucleotides. b, Alignment of DZIP1 and DZIP2 protein sequences indicated a >80% identical amino acids shared. Black shaded area represents identical amino acids, and gray boxed area outlines the C2H2 zinc-finger domain. GenBank accession nos.: DZIP1, AF272347; DZIP2, AF272348.

Figure 2 Expression of *DZIP1* and *DZIP2* mRNAs. a, Northern blot analysis of adult tissues indicated *DZIP1* is predominantly expressed in testis. Three *DZIP1* transcripts of approximately 1.35, 3.9, and 4.4 kilobases are detected. b, Northern blot analysis of adult tissues indicated *DZIP2* is testis specific with a transcript of approximately 3.9 kilobases. c, RT-PCR expression analysis of *DZIP1* and *DZIP2* in embryonic stem (ES) cells, fetal tissues, and adult tissues detected early expression. *DZIP1* transcripts are detected in ES cells, fetal ovary and testis, and adult testis. *DZIP2* transcripts are detected in ES cells, fetal brain, fetal ovary, fetal testis, and adult testis.

Figure 3 Specific interactions of DZIP1 and DZIP2 and DAZ. a, The DAZ protein coimmunoprecipitated with DZIP1 when HA beads were incubated with yeast supernatant from cells expressing both DAZ and DZIP1 fused to an HA tag; shown is DAZ protein detected by Western analysis (left panel). DAZ protein was not

coimmunoprecipitated with HA beads in the presence of yeast expressing DAZ or DZIP1 alone. Right lane shows supernatant from yeast expressing DAZ as a positive control. Right panel: Likewise, DAZ protein was coimmunoprecipitated with DZIP2 when incubated with HA beads in yeast supernatant from cells expressing both DAZ and DZIP2 fused to an HA tag. The same controls were used as described above. b, Deletions of DAZ protein define the minimal region of DAZ required for interaction with DZIP2. The yeast two-hybrid assay was used to test the ability of eight deletions of the DAZ:GAL4 fusion protein to interact with DZIP2. Interactions were assayed using the X-GAL filter assay (according to manufacturer's protocol; Clontech, Inc., Palo Alto, CA). c, Deletions of the DZIP2 protein define the minimal domain of DZIP2 required for interaction with DAZ. Five deletions of the DZIP2:GAL4 fusion protein were constructed and assayed for their ability to interact with intact DAZ protein in the yeast two-hybrid system using an X-GAL filter assay. Assays were done as indicated in Methods. +++, maximal activity equivalent to that of intact interacting proteins; ++, nearly wild-type interaction; +, strong interaction; +/–, weak interaction; –, no interaction.

Figure 4 DZIP1 localizes to specific compartments in fetal and adult germ cells. a, Western blot analysis with DZIP1 antisera (lanes 1-3) and preimmune antisera (lanes 4-6). Cell extracts are from human testis (lane 1,4,), mouse testis (lane 2,5), and mouse ES cells (lane 3,6). The major DZIP1 protein bands in human testis are ~ 89 kD and 123 kD. Protein bands of ~80 kD and 120 kD were observed in extracts from mouse testis, and ~92 kD and 123 kD protein bands were detected in mouse ES cells. b, Expression of *DZIP2* is detected in men with germ cells (normal testis) and not detected in men without



germ cell (SCO). c, DZIP1 protein is not detected in fetal testis with pre-immune antisera. d, Localization of DZIP1 in fetal testis is detected in the nucleus and cytoplasm of PGCs. e, DZIP1 is not detected in adult testis with pre-immune antisera. f, In adult testis DZIP1 is detected in spermatogonia, primary and secondary spermatocytes. g, DZIP1 is not detected in fetal ovaries with pre-immune antisera. h, DZIP1 is also detected in the PGCs in fetal ovaries. Magnification: X400. (SCO) sertoli cell only; (Spg) spermatogonia; (pSpc) primary spermatocyte; (Spc) spermatocyte

Figure 5 Cellular and subcellular localization of DZIP1 in mouse testis. a, DZIP1 localizes in adult mouse testis to the nucleus and cytoplasm in spermatogonia and to the cytoplasm in spermatocytes. b, Immunofluresence staining of DZIP1 in mouse testis. c, DZIP1 protein is not detected with pre-immune antisera in mouse testis. d, Subcellular staining of DZIP1 in early round spermatids indicates increase localization in proacrosome vesicle. e, DZIP1 staining is detected in the acrosome of round spermatids. f, In elongating spermatids, DZIP1 is also detected in the cytoplasm and acrosomal head.

References

1. Saffman, E.E. and P. Lasko, *Germline development in vertebrates and invertebrates.* Cell. Mol. Life. Sci., 1999. **55**: p. 1141-63.

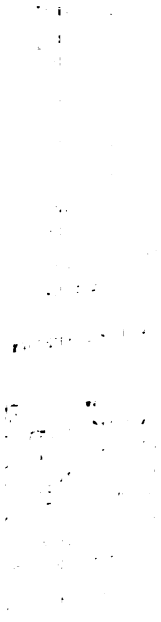
2. Reijo, R., et al., *Diverse spermatogenic defects in humans caused by Y chromosome deletions encompassing a novel RNA-binding protein gene.* Nat. Genet., 1995. **10**: p. 383-93.

3. Ruggiu, M., et al., *The mouse Dazla gene encodes a cytoplasmic protein essential for gametogenesis.* Nature, 1997. **389**: p. 73-7.

4. Houston, D.W. and M.L. King, *A critical role for Xdazl, a germ plasm-localized RNA, in the differentiation of primordial germ cells in Xenopus.* Devel., 2000. **127**: p. 447-56.

5. Dorfman, D.M., D.R. Genest, and R.A.R. Pera, *Human DAZL1 encodes a candidate fertility factor in women that localizes to the prenatal and postnatal germ cells.* Hum. Reprod., 1999. **14**: p. 2531-2536.

6. Reijo, R.A., et al., *DAZ family proteins exist throughout male germ cell development and transit from nucleus to cytoplasm at meiosis in humans and mice.* Biol. Reprod., 2000. **63**: p. 1490-1496.



7. Johnson AD, B.R., Drum M, Masi T., *Expression of axolotl dazl RNA, a marker of germ plasm: widespread maternal RNA and onset of expression in germ cells approaching the gonad.* Dev. Biol., 2001. **234**: p. 402-15.

8. Maegawa, S., K. Yasuda, and K. Inoue, *Maternal mRNA localization of zebrafish DAZ-like gene.* Mech. Dev., 1999. **81**: p. 223-26.

9. Moore, F.L., Jaruzelska, J., Fox, M.S., Urano, J., Firpo, M.T., Turek, P.J., Dorfman, D.M., Reijo-Pera, R.A., *A Human Homolog of Drosophila Pumilio is Expressed in Embryonic Stem Cells and Germ Cells and Interacts with DAZ (Deleted in AZoospermia) and DAZ-Like Proteins.* Submitted to PNAS, 2002.

10. Lin, H. and A.C. Spradling, *A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary.* Development, 1997. **124**: p. 2463-76.

11. Foresta, C., et al., *Y-chromosome deletions in idiopathic severe testiculopathies.* J. Clin. Endocrinol. Metab., 1997. **82**: p. 1075-80.

12. Xu, E.Y., F.L. Moore, and R.A.R. Pera, *A gene family required for human germ cell development evolved from an ancient meiotic gene conserved in all metazoans.* Proc. Natl. Acad. Sci., 2001. **98**: p. 7414-19.

13. Cheng, M.H., J.Z. Maines, and S.A. Wasserman, *Biphasic subcellular localization of the DAZL-related protein Boule in Drosophila spermatogenesis.* Dev. Biol., 1998. **204**: p. 567-576.

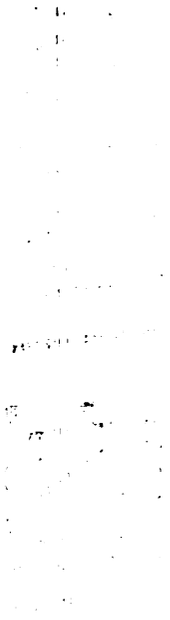
14. Eberhart, C.G., J.Z. Maines, and S.A. Wasserman, *Meiotic cell cycle requirement for a fly homologue of human Deleted in AZoospermia.* Nature, 1996. **381**: p. 783-5.

15. Maines, J.Z. and S.A. Wasserman, *Post-transcriptional regulation of the meiotic Cdc25 protein Twine by the Dazl orthologue Boule.* Nature Cell. Biol., 1999. **1**: p. 171-174.

16. Yen, P.H., N.N. Chai, and E.C. Salido, *The human autosomal gene DAZLA: testis specificity and a candidate for male infertility.* Hum. Mol. Genet., 1996. **5**: p. 2013-7.

17. Shan, Z., et al., *A SPGY copy homologous to the mouse gene Dazla and the Drosophila gene boule is autosomal and expressed only in the human male gonad.* Hum. Mol. Genet., 1996. **5**: p. 2005-11.

18. Cooke, H.J., et al., *A murine homologue of the human DAZ gene is autosomal and expressed only in male and female gonads.* Hum. Mol. Genet., 1996. **5**: p. 513-6.



19. Saxena, R., et al., *Four DAZ genes in two clusters found in the AZFc region of the human Y chromosome.* Genomics, 2000. **67**: p. 256-67.

20. Carani, C., et al., *cynDAZLA: a cynomolgus monkey homologue of the human autosomal DAZ gene.* Mol. Hum. Reprod., 1997. **3**: p. 479-483.

21. Saxena, R., et al., *The DAZ gene cluster on the human Y chromosome arose from an autosomal gene that was transposed, repeatedly amplified and pruned.* Nat. Genet., 1996. **14**: p. 292-9.

22. Gromoll, J., et al., *The Old World monkey DAZ (Deleted in AZoospermia) gene yields insights into the evolution of the DAZ gene cluster on the human Y chromosome.* Hum. Mol. Genet., 1999. **8**: p. 2017-24.

23. Slee, R., et al., *A human DAZ transgene confers partial rescue of the mouse Dazl null phenotype.* Proc. Natl. Acad. Sci., 1999. **96**: p. 8040-45.

24. Houston, D.W., et al., *A Xenopus DAZ-like gene encodes an RNA component of germ plasm and is a functional homologue of Drosophila boule.* Development, 1998. **125**: p. 171-80.

25. Wickens, M., et al., *A PUF family portrait: 3'UTR regulation as a way of life.* Trends Genet. 2002. **18**: p. 150-7.

26. Parisi, M. and H. Lin, *The Drosophila Pumilio gene encodes two functional protein isoforms that play multiple roles in germline development, gonadogenesis, oogenesis, and embryogenesis.* Genetics, 1999. **153**: p. 235-50.

27. Lehmann, R. and C. Nusslen-Volhard, *Involvement of the Pumilio gene in the transport of an abdominal signal in the Drosophila embryo.* Nature, 1987. **329**: p. 167-170.

28. Murata, Y. and R.P. Wharton, *Binding of Pumilio to Maternal hunchback mRNA Is Required for Posterior Patterning in Drosophila Embryos.* Cell, 1995. **80**: p. 747-756.

29. Forbes, A. and R. Lehmann, *Nanos and Pumilio have critical roles in the development and function of Drosophila germline stem cells.* Devel., 1998. **125**: p. 679-690.

30. Caffrey, C.R., et al., *SmCB2, a novel tegumental cathepsin B from adult Schistosoma mansoni.* Mol. Biochem. Parasitol., 2002. **121**: p. 49-61.

31. Habermann, B., et al., *DAZ (Deleted in AZoospermia) genes encode proteins located in human late spermatids and in sperm tails.* Human Reproduction, 1998. **13**: p. 363-369.

32. Lin, Y.M., et al., *Expression patterns and transcript concentrations of the autosomal DAZL gene in testes of azoospermic men.* Mol. Hum. Reprod., 2001. **7**: p. 1015-22.

33. Sonoda, J. and R.P. Wharton, *Drosophila Brain Tumor is a translational repressor.* Genes Dev., 2001. **15**: p. 762-73.

34. Sonoda, J. and R.P. Wharton, *Recruitment of Nanos to hunchback mRNA by Pumilio.* Genes Dev., 1999. **13**: p. 2704-12.

35. Ruggiu, M. and H. Cooke, *In vivo and in vitro analysis of homodimerization activity of the mouse Dazl protein.* Gene, 2000. **252**: p. 119-126.

36. Tsui, S., et al., *Identification of two novel proteins that interact with germ-cell specific RNA-binding proteins DAZ and DAZL1.* Genomics, 2000. **65**: p. 266-73.

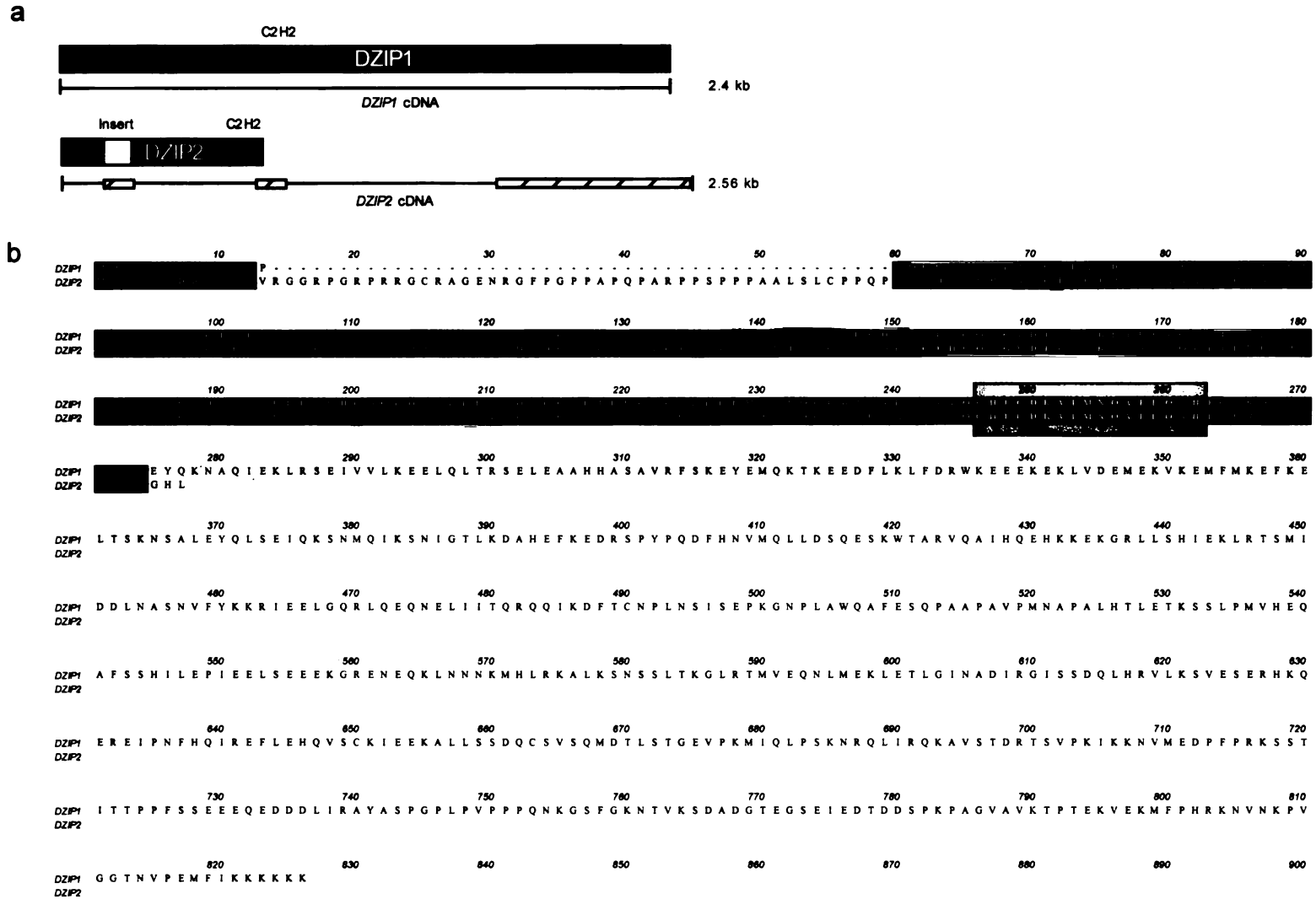

a
C2H2
DZIP1
2.4 kb
DZIP1 cDNA
Insert
C2H2
DZIP2
2.56 kb
DZIP2 cDNA
b
DZIP1
DZIP2

Figure 1

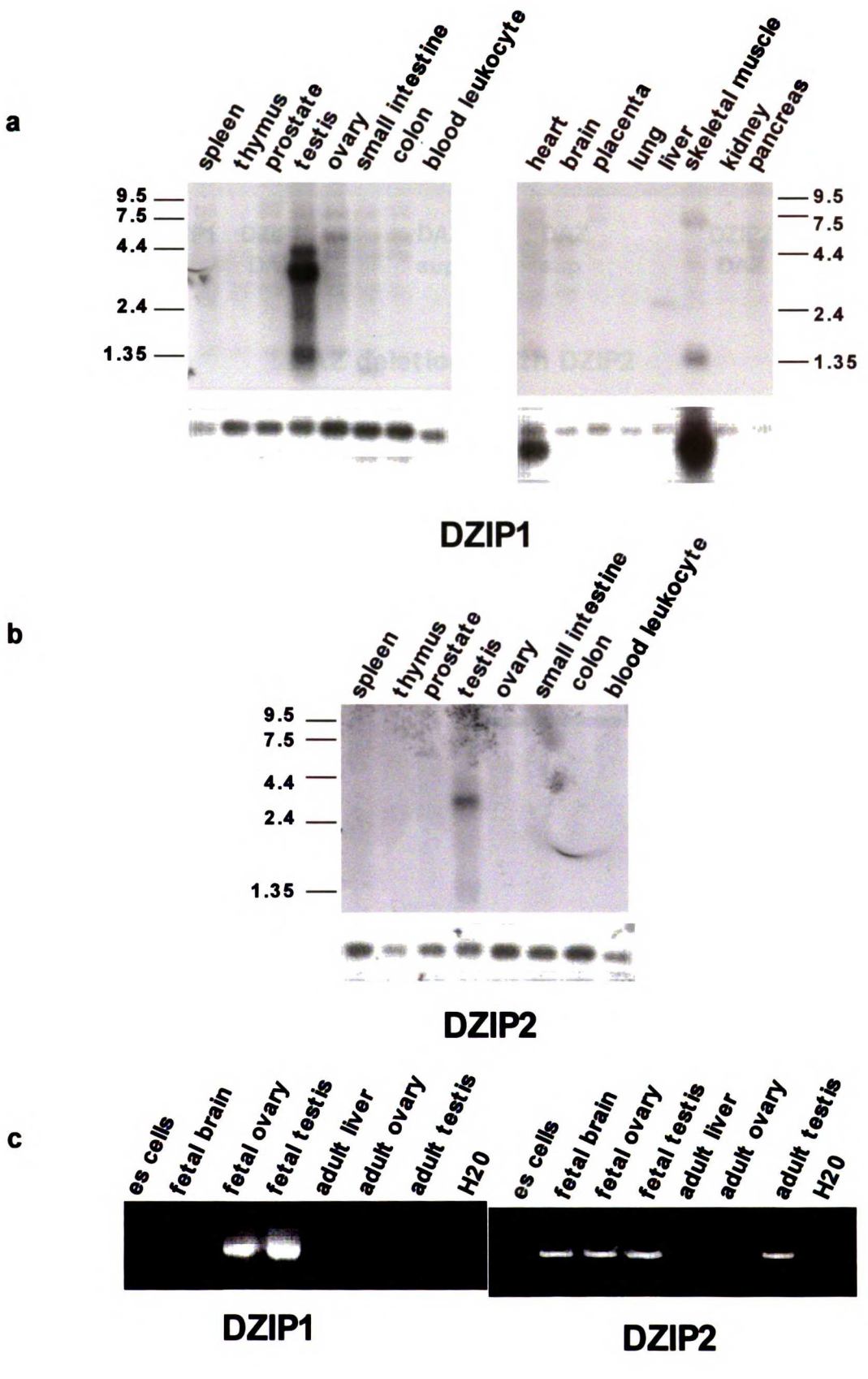
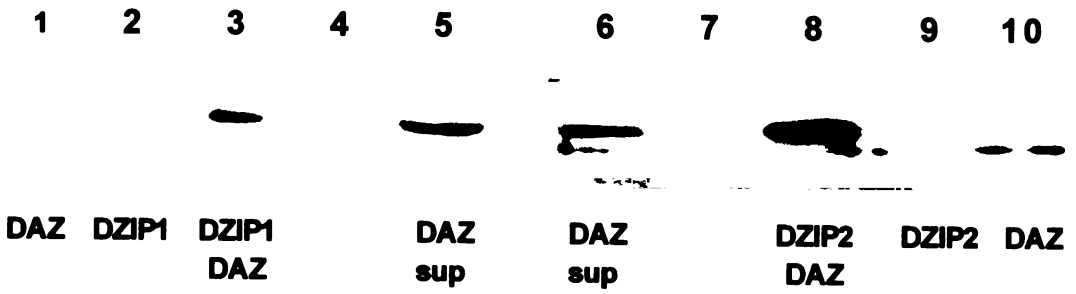
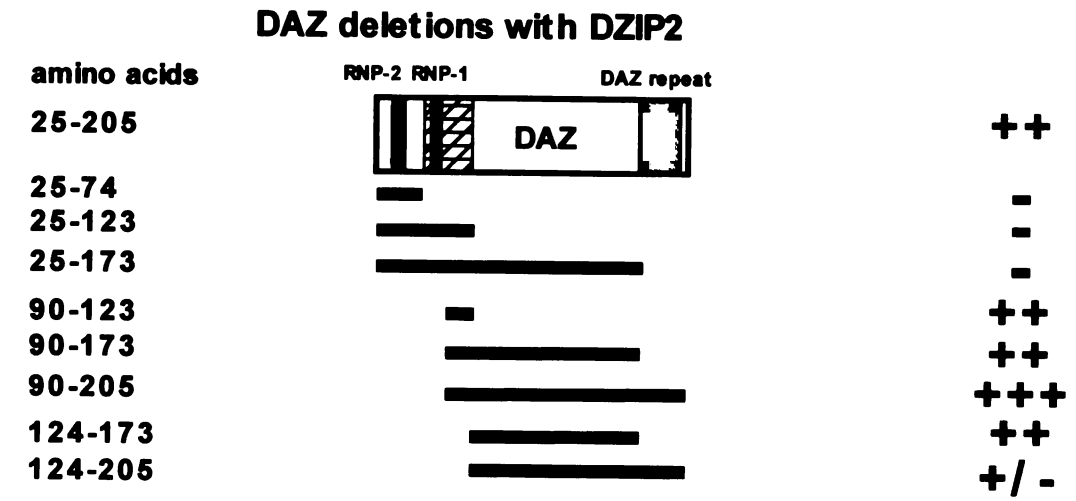

a
spleen
thymus
prostate
testis
ovary
small intestine
colon
blood leukocyte
heart
brain
placenta
lung
liver
skeletal muscle
kidney
pancreas
9.5
7.5
4.4
2.4
1.35
DZIP1
b
spleen
thymus
prostate
testis
ovary
small intestine
colon
blood leukocyte
9.5
7.5
4.4
2.4
1.35
DZIP2
c
es cells
fetal brain
fetal ovary
fetal testis
adult liver
adult ovary
adult testis
H20
DZIP1
DZIP2

Figure 2

a


1 2 3 4 5 6 7 8 9 10
DAZ DZIP1 DZIP1 DAZ DAZ sup DAZ sup DZIP2 DAZ DZIP2 DAZ

b


DAZ deletions with DZIP2
amino acids
RNP-2 RNP-1
DAZ repeat
DAZ
25-205 ++
25-74 -
25-123 -
25-173 -
90-123 ++
90-173 ++
90-205 +++
124-173 ++
124-205 +/-

c

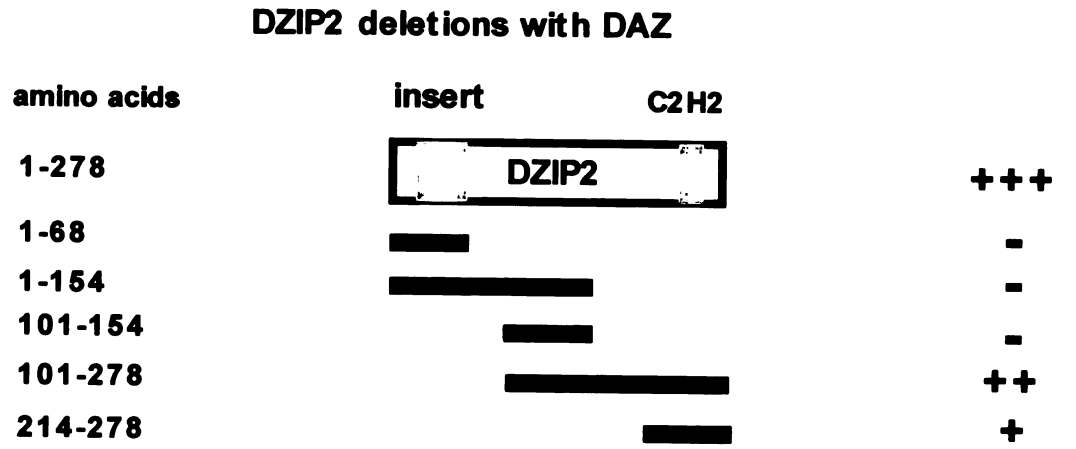

DZIP2 deletions with DAZ
amino acids
insert
C2H2
DZIP2
1-278 +++
1-68 -
1-154 -
101-154 -
101-278 ++
214-278 +

Figure 3

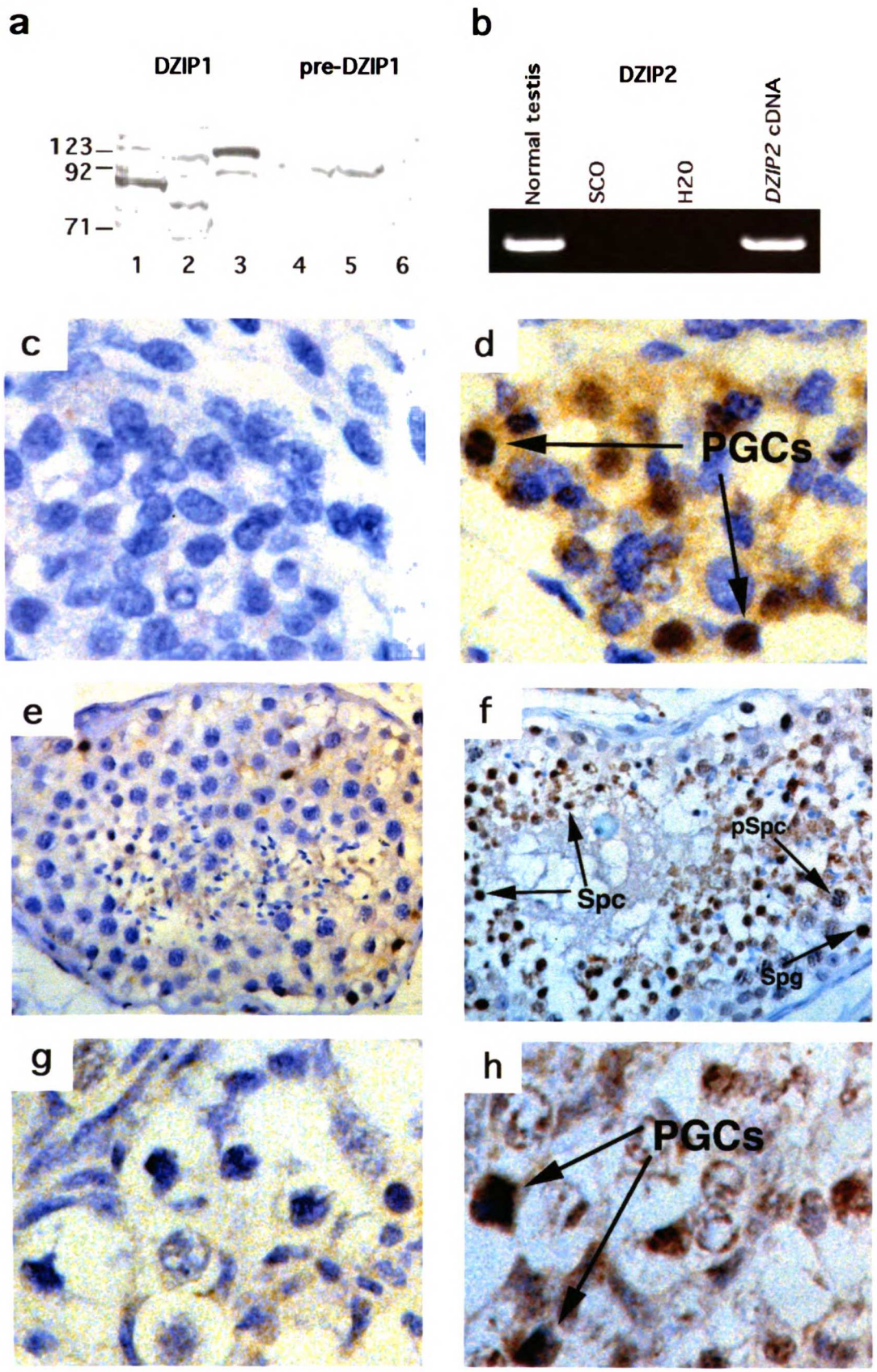
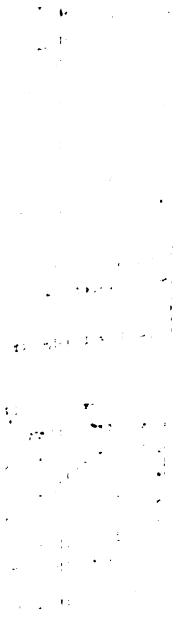

a
DZIP1
pre-DZIP1
123
92
71
1 2 3 4 5 6
b
DZIP2
Normal testis
SCO
H20
DZIP2 cDNA
c
d
PGCs
e
f
pSpc
Spc
Spg
g
h
PGCs

Figure 4



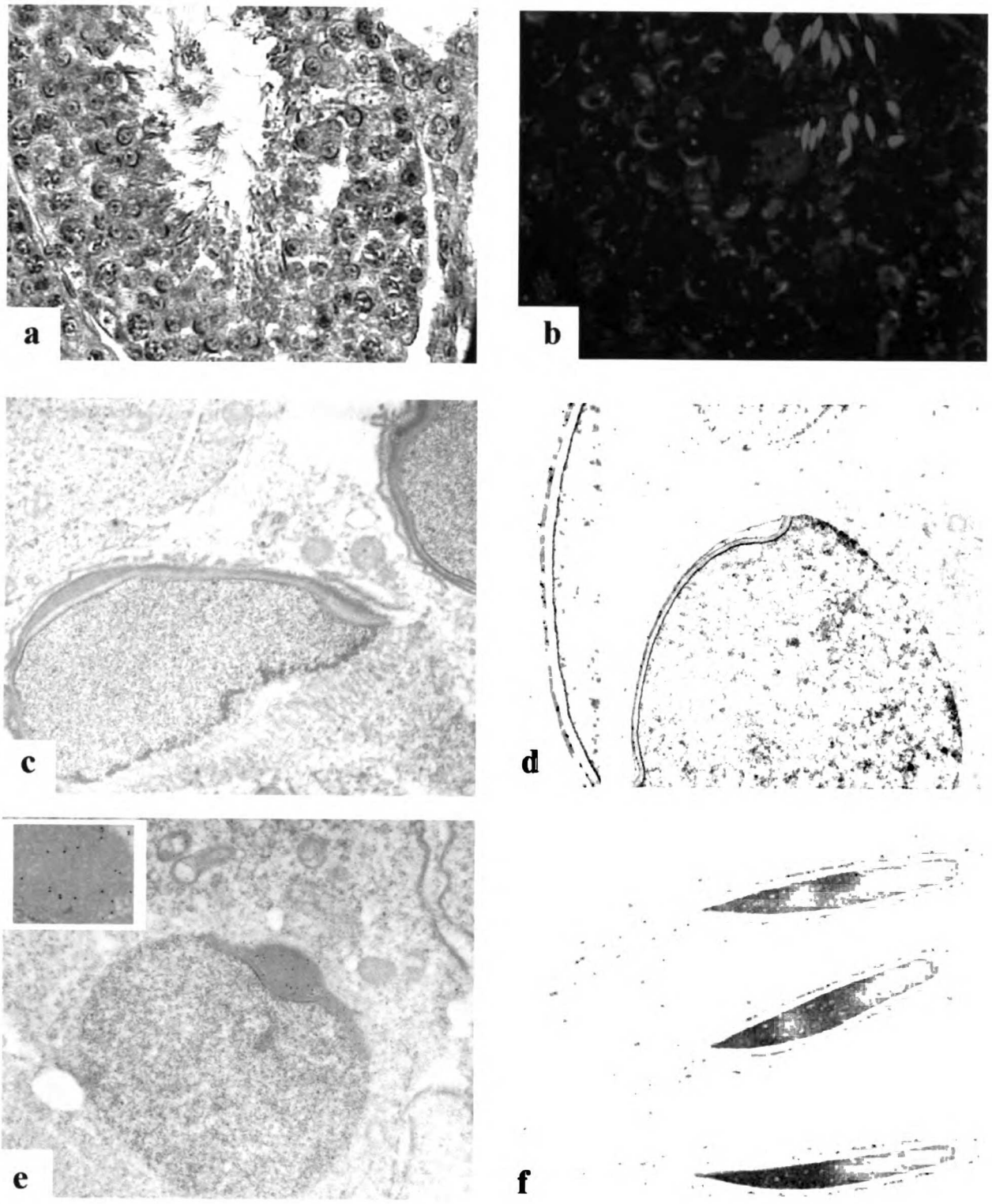
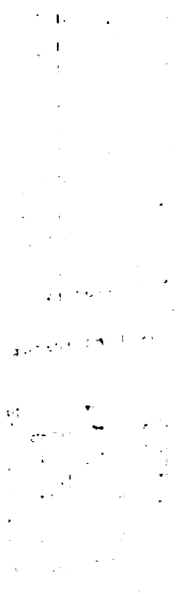


Figure 5



Chapter 5

The Human *Quaking* Gene Codes for a Novel DAZ-Interacting Protein

The Human *Quaking* Gene Codes for a Novel DAZ-Interacting Protein

Frederick L. Moore* and Renee A. Reijo-Pera*

*Department of Obstetrics, Gynecology and Reproductive Sciences; Department of Physiology; Department of Urology; and Programs in Human Genetics and Cancer Genetics; University of California at San Francisco, San Francisco, CA 94143-0546, USA

Abstract

The two-hybrid screen that was used to identify DAZ-interacting proteins uncovered a novel cDNA transcribed from the human *Quaking* locus. A spontaneous deletion that removes part of the 5' end of the *quaking* locus in the mouse causes a pleiotrophic phenotype, one of which is sterility in the male mouse[1]. We characterized the expression of the human *Quaking 3* (*QK3*) transcript and verify that the protein that it encodes forms a stable complex with the DAZ protein by co-immunoprecipitation in yeast cells. We examine the germ cells in the testis of *quakingviable* (qk^v) mice at different stages in germ cell development and verified previous work that addressed the cause of sterility due to arrest of spermatid maturation[2]. Based on information from model organisms, we propose a genetic model for the cause of sterility in the qk^v mouse and propose a molecular model for the function of QK3 and QKI proteins in human and mouse germ cell development.

Results & Discussion

Identification of the human *Quaking 3* (*Qk3*) gene

In our two-hybrid screen for DAZ-interacting proteins, we isolated one clone that represented a cDNA transcribed from the *quaking* locus. Originally characterized in the mouse, the *quaking* (*quaking*^{viable}, *qk*^{v}) phenotype results from a spontaneous recessive mutation that causes a severe deficit in myelin[5]. As a consequence, *qk*^{v} homozygotes develop a rapid tremor by postnatal day 10. Interestingly, the *qk*^{v} mutation also results in sterility of homozygous males due to the arrest of spermatid maturation[1]. While the original *quaking* mutation is viable, paradoxically, five recessive, ethylnitrosourea (ENU)- induced alleles (collectively referred as *qk*^{e}) are embryonic lethal[6-9]. The cloning of a candidate gene, *qkI,* which encodes five protein isoforms, QKI-5a, -5b, -6, -7, and –G, which are abundant in myelin-forming cells of wild type mice, but whose levels are severely reduced in myelin-forming cells of qk^{v} mice, helped to further elucidate a genetic cause for the *quaking* phenotype[10]. It was determined that the ENU mutations that caused lethality in the *qk*^{e} mice created amino acid changes in the coding region of the *qkI* transcripts, which disrupted the QKI function[8, 11-13]. Additional results from the cloning of *qkI* indicated that the spontaneous mutation in the qk^{v} mouse was a deletion in the 5' end of the *quaking* locus. The *qkI* transcripts were not affected by the deletion in the qk^{v} mouse, yet a previously uncharacterized transcript called *qkII*, which is located in the region deleted in the qk^{v} mouse was affected[10] (Fig. 1a). The open reading frame for the *qkII* transcript and its function is unknown. Interestingly, the *QK3* transcript isolated from our two-hybrid screen contains part of the 3' end of the *qkII*

transcript and a portion of the 5' end of the *qkI* transcript (Fig. 1a). Also of interest, the theoretical protein translated from the *QK3* cDNA when aligned to a theoretical protein translated from the *qkII* cDNA shows greater than 64% homology (Fig. 1b). There are no apparent domains contained within the QK3 protein sequence.

Expression of QK3 transcript

We next sought to characterize the expression of *QK3* in human tissue. Northern analysis using a probe that covered a small region unique to the *qkII* transcript, yet also contained a large portion of the mouse *qkI* transcripts, detected a prominent band of ~4.0 kilobases, highly expressed in the testis and ovaries (Fig. 2a). A second band of ~4.6 kilobases was detected in multiple tissues, including the testis and ovaries (Fig. 2a, b). An additional experiment that utilizes a probe that contains only the region similar to the mouse *qkII* transcript (first 300 bps) would give additional information on the specific expression of the novel *QK3* transcript.

Verification of QK3 interaction with DAZ protein

We next sought to further verify that the DAZ protein interacts with QK3 protein by co-immunoprecipitation in yeast cells. Yeast supernatant was derived from yeast expressing DAZ alone, QK3 alone, and both DAZ and QK3. We utilized an HA-tag fused to the N-terminus of QK3 to isolate it from yeast supernatant and probed using Western analysis for the presence of DAZ protein. Our results show that the DAZ protein is only isolated from the yeast supernatant containing both DAZ and QK3 proteins (Fig.

2b, lane 3). These data shows that the DAZ protein can form a stable complex with QK3 protein.

Localization of QK3 protein & analysis of the *quaking* mouse testis

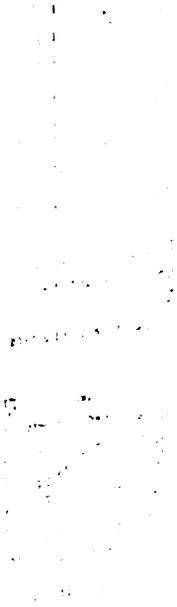
To determine the subcellular localization of the QK3 protein in human and mouse germ cells, a peptide specific to the QK3 protein was designed to generate polyclonal antibodies in rabbit. Unfortunately, the antisera against the QK3 peptide were nonspecific, as judged by Elisa experiments. In parallel, we ordered the mutant qk^{v} mouse, which is available commercially (Jackson Research Laboratory). Our original interest was to determine the localization of QK3 protein in the quaking (qk^{v}) mouse compared to the wild type mouse. We were also interested in determining the localization of DAZ and DAZ-associated proteins in the qk^{v} mouse. I believe that by understanding the normal localization of the QK3 protein in germ cells, in conjunction with DAZ and DAZ-associated proteins, will assist in building a functional model for the role of QK3 and its interacting partners.

Previous work has determined that in the qk^{v} male mouse there is a decrease in the number of spermatids compared to normal, due to an arrest in spermatid maturation[2]. Chubb at el. determined that the germ cell number in the qk^{v} mouse is normal through the pachytene stage of meiosis, but at specific stages of spermiogenesis the germ cell numbers are lowered. We decided to utilize the qk^{v} mice available to us by collecting testis tissue from different stages (9, 15, & 21 day postnatal) of germ cell development to verify that the male mice produce sperm in the adult testis and that the

apparent number of germ cells at the onset of spermiogenesis (21 day) is comparable in qk^v mice versus normal. We show in figure 3b that spermatids are present in the seminiferous tubule of testis sections from adult qk^v mice. In addition, the number of germ cells present in the seminiferous tubules of qk^v mice compared to normal mice at 21 day postnatal is comparable (Fig. 3c, d).

Model for sterility of qk^v male mice

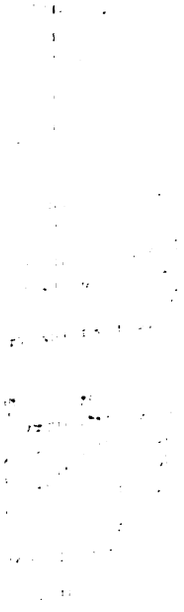
The spontaneous recessive mutation that results in the qk^v mouse is caused by a 1 Mb deletion that disrupts the 5' end of the *quaking* locus. Although the qk^v male mice are sterile due to arrest of spermatid maturation[1], it is now thought that the sterility is due to a separate gene deleted in the qk^v mouse[14]. There are five ENU point mutations, which disrupt various regions of the open reading frame of the *qkI* gene, yet when these alleles are heterozygous with qk^v, quaking but fertile males are obtained[10]. Homozygosity of the ENU-induced alleles results in embryonic lethality[11, 15]. One possible explanation for the sterility in the qk^v mouse is the disruption of the *qkII* transcript. We have demonstrated that the QK3 protein, which is translated from the *QK3* cDNA that contains regions of both, *qkI* and *qkII* transcripts, makes a protein that can form a stable complex with the DAZ protein. Also, the *qkII* transcript, which has an open reading frame that is unknown, can theoretically translate a protein that is a homolog to QK3 protein. Additional experiments can elucidate if *qkII* is the cause of sterility in the qk^v mouse. For example, the introduction of the *qkII* transgene into the qk^v mice would address if *qkII* can rescue the sterility in the qk^v male mice.



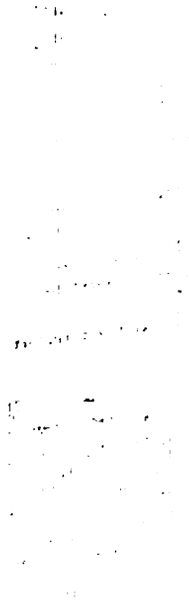
Molecular Model of QK3 Function with the DAZ Protein

There are various possible models for the function of QK3 (qkII) protein and QKI proteins in human and mouse germ cell development. I will speculate on possible models based on information obtained on *quaking* homologs from model organisms. The mouse *quaking* (*qk*) gene expresses at least six different alternatively spliced mRNA including QKI-5, QKI-6, QKI-7, and QKI-G, which differ in their C-terminal 30 amino acids[10, 13, 16]. The KH domain of the QKI proteins is embedded in a larger conserved domain of ~200 amino acids, called the STAR (signal transduction activator of RNA metabolism) domain[17]. The STAR domain of the QKI proteins is required for RNA binding and dimerization[12, 18-20]. The cellular localization of the QKI isoforms differ, and evidence suggest that sequences in the C-terminus play specific roles in their localization and function. In the mouse, the QKI-5 protein contains a motif at the C-terminus, called STAR-NLS, which localizes the protein to the nucleus[20]. The QKI-5 protein has been shown to shuttle between the nucleus and cytoplasm[20]. The QKI-7 protein in the mouse also contains a unique sequence at the C-terminus that functions to induce apoptosis[21]. The normal localization of QKI-7 is to the cytoplasm, and hetero-dimerization of QKI-5 and QKI-7 localizes QKI-7 to the nucleus and inactivates apoptosis[21]. In addition, the QKI-6 protein in the mouse has been shown to localize to the cytoplasm with low levels in the nucleus[22], and evidence indicates that QKI-6 can act as a translational repressor in vitro and in vivo[23].

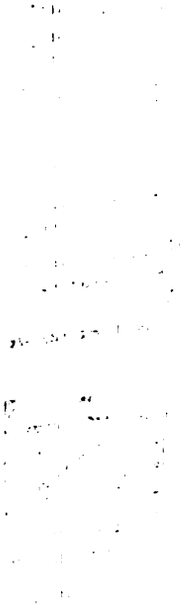
If the interaction between DAZ and QK3 proteins is legitimate, then it raises the question: Are there other interactions between the DAZ family of proteins and the



Quaking family of proteins? If the DAZ and DAZL proteins interact with the QKI proteins, this interaction allows insight into possible mechanisms for the localization of proteins, post-transcriptional regulation of transcripts, and the regulation of apoptosis within the germ cell development pathway. One hypothesis is that DAZ and DAZL proteins are sequestered to the nucleus in spermatogonia cells through interaction with QKI-5 and relocated to the cytoplasm at the onset of meiosis through interactions with additional proteins (Fig. 4). The QKI-5 protein in the mouse has been shown to regulate the localization of QKI-7 from the cytoplasm to the nucleus by dimerization[21]. Another hypothesis for a molecular function for QK3 and DAZ is competition between QK3 and the QKI proteins. For example, SAM68, a member of the STAR family of proteins that binds RNA and plays a suggested role in signal transduction[24], has a rare alternative splice form lacking most of the KH domain. Interestingly, it has a vastly lower RNA-binding affinity and is specifically found in growth-arrested cells, where it seems to antagonize full-length SAM68 in cell-cycle progression[25]. A connection between the DAZ family of proteins and the STAR family of proteins is tempting because it connects new information being discovered in other model organisms. GLD-1, a member of the STAR protein family is the closest relative to QKI. In *C. elegans*, *gld-1* has an essential function in oocyte differentiation and meiotic prophase progression. The FBF protein, which is a homolog of the *Drosophila* and human Pumilio proteins, regulates germline stem cell maintenance by repressing translation of the *GLD-1* transcript[26]. Our previous work has shown that the DAZ and DAZL proteins interact with PUM2 in humans. Future experiments will help to determine if PUM2 and/or DAZ protein(s) regulate the translation of *quaking* transcripts and/or members in the STAR



family and provide us with a better framework to understand the details of the germ cell developmental pathway in humans and mouse.



Materials & Methods

Identification of *QK3*

The yeast two-hybrid system was used to screen for proteins that interact with DAZ protein expressed as a fusion with the GAL4 DNA-binding domain (Clontech, Inc., Palo Alto, CA). This construct was derived from a cDNA that contains the N-terminal RNA-binding domain and a single DAZ repeat (pRR102). A testis library of random cDNAs fused to the activation domain of GAL4 was transformed into yeast and screened according to the manufacturer's instructions. *QK3* cDNA was isolated from our screen once. The encoded protein for QK3 interacted specifically with DAZ protein and did not interact with the DNA-binding domain alone, DNA-binding domain fused to laminin, or activated transcription independent of the DAZ fusion protein. Sequencing of the *QK3* clone was done at the Genome Core Facility of the Program in Human Genetics (UCSF).

Northern Analysis

Restriction enzymes NarI and EcoRV, which cut at sites 105 and 1949 on the *QK3* plasmid respectively, were used to isolate a cDNA probe for Northern analysis. The probed was radiolabeled using a random primer labeling kit (ROCHE) and hybridized to a polyadenylated RNA blot (CLONTECH).

Co-immunoprecipitation of QK3 & DAZ in Yeast Cells

Co-immunoprecipitation was done in yeast extracts. Yeast cells expressing DAZ, QK3, and both DAZ and QK3 were pelleted, lysed, and resuspended in lysis buffer (50 mM Tris-HCl, pH 7.5, 1 mM EDTA, 150 mM NaCl, 10% glycerol, 1% NP-40) with a

cocktail of protease inhibitors (Clontech, Inc.). The *QK3* cDNA, which is fused to the GAL4 activation domain and contains a hemagglutinin (HA) tag, was used to immunoprecipitate QK3 from yeast supernatant. HA antibodies covalently linked to inert beads (BABCO, Inc.) were incubated with yeast supernatants, washed with the lysis buffer 3 times, resuspended in 1 mM glycine (pH 2.8), and incubated at 30°C to elute bound proteins. Proteins were separated on a sodium dodecyl sulfate-polyacrylamide gel. The presence of DAZ proteins was assessed by Western analysis with DAZ-specific antibodies [3, 4].

Staining of Quaking Mouse Testis

To determine the localization of the QK3 protein in human and mouse testis tissue, a peptide with sequence "AQAAAALSGLDQPSERRPRL" was designed to generate polyclonal antibodies in rabbits. The Elisa reading for the antibody serum indicated nonspecific binding to the chosen peptide. Hematoxylin/Eosin (H/E) staining was carried out with sections of 21 day and adult testis tissue, from heterozygote and homozygote *quaking* mice, according to protocol (Current Protocols in Molecular Biology).

Figure legend

Figure 1 Analysis of *QK3* cDNA and protein sequence. a, A schematic of the mouse *qkI* and *qkII* transcripts aligned with the *QK3* cDNA from humans. *QK3* shares regions of both *qkI* and *qkII* transcripts. A spontaneous deletion in the 5' UTR region of *qkI* results in male sterility. The proteins encoded by *qkI* contain a KH RNA-binding domain and QUA1 and QUA2 domains. The protein encoded by the human *QK3* transcript is novel and contains no known domains. b, Alignment of the QK3 protein compared to a theoretical protein translated from the *qkII* cDNA shows greater than 64% homolog.

Figure 2 Expression of *QK3* mRNA and interaction of DAZ and QK3 proteins. a,b, Northern analysis of *QK3* transcript in human tissue detects a predominant band of ~4.2 kb in testis and ovaries. An additional band of ~4.6 kb is detected in multiple tissues. c, The DAZ protein coimmunoprecipitated with QK3 when HA beads were incubated with yeast supernatant from cells expressing both DAZ and QK3 fused to an HA tag; shown is DAZ protein detected by Western analysis. DAZ protein was not coimmunoprecipitated with HA beads in the presence of yeast expressing DAZ or QK3 alone. Right lane shows supernatant from yeast expressing DAZ as a positive control.

Figure 3 Analysis of *quaking* versus wild-type mice testis tissue. a, Adult heterozygous *quaking* mice are fertile and exhibit normal spermatogenesis. b, Adult homozygous *quaking* mice are sterile, yet exhibit germ cells in their testis that have completed meiosis and matured into round spermatids. c,d, 21 day heterozygous and homozygous *quaking*

mice exhibit spermatocytes in their testis tubules that are progressing into spermiogenesis.

Figure 4 Model of Quaking proteins function in the testis. If QKI-5 protein, which is a member of the Quaking family of proteins, binds to the DAZ and DAZL proteins, then QKI-5 may function to shuttle DAZ and DAZL from the nucleus to the cytoplasm at the onset of meiosis to regulate function. QK3 protein may interact with DAZ protein to compete for its binding to QKI proteins in the cytoplasm of germ cells during different times of development.

References

1. Bennett, W.I., et al., *Abnormal spermiogenesis in quaking, a myelin-deficient mutant mouse.* Biol. Reprod., 1971. **5**: p. 30-58.

2. Chubb, C., *Oligotriche and quaking gene mutations. Phenotypic effects on mouse spermatogenesis and testicular steroidogenesis.* J. Androl., 1992. **13**: p. 312-7.

3. Dorfman, D.M., D.R. Genest, and R.A.R. Pera, *Human DAZL1 encodes a candidate fertility factor in women that localizes to the prenatal and postnatal germ cells.* Hum. Reprod., 1999. **14**: p. 2531-2536.

4. Reijo, R.A., et al., *DAZ family proteins exist throughout male germ cell development and transit from nucleus to cytoplasm at meiosis in humans and mice.* Biol. Reprod., 2000. **63**: p. 1490-1496.

5. Sidman, R.L., Dickie M.M., Appel S.H., *Mutant mice (quaking and jimpy) wirth deficeint myelination in the central nervous system.* Science, 1964. **144**: p. 309-312.

6. Bode, V.C., *Ethylnitrosourea mutagenesis and the isolation of mutant alleles for specific genes located in the T region of mouse chromosome 17.* Genetics, 1984. **108**: p. 457-70.

7. Justice, M.J. and V.C. Bode, *Induction of new mutations in a mouse t-haplotype using ethylnitrosourea mutagenesis.* Genet. Res., 1986. **47**: p. 187-92.

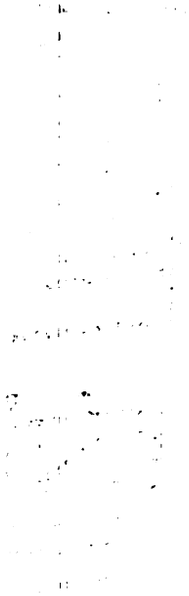
8. Shedlovsky, A., T.R. King, and W.F. Dove, *Saturation germ line mutagenesis of the murine t region including a lethal allele at the quaking locus.* Proc. Natl. Acad. Sci. U S A, 1988. **85**: p. 180-4.

9. Justice, M.J. and V.C. Bode, *Genetic analysis of mouse t haplotypes using mutations induced by ethylnitrosourea mutagenesis: the order of T and qk is inverted in t mutants.* Genetics, 1988. **120**: p. 533-43.

10. Ebersole, T.A., et al., *The quaking gene product necessary in embryogenesis and myelination combines features of RNA binding and signal transduction proteins.* Nat. Genet., 1996. **12**: p. 260-5.

11. Justice, M.J. and V.C. Bode, *Three ENU-induced alleles of the murine quaking locus are recessive embryonic lethal mutations.* Genet. Res., 1988. **51**: p. 95-102.

12. Chen, T. and S. Richard, *Structure-function analysis of Qk1: a lethal point mutation in mouse quaking prevents homodimerization.* Mol. Cell Biol., 1998. **18**: p. 4863-71.



13. Cox, R.D., et al., *Contrasting effects of ENU induced embryonic lethal mutations of the quaking gene.* Genomics, 1999. **57**: p. 333-41.

14. Ebersole, T., O. Rho, and K. Artzt, *The proximal end of mouse chromosome 17: new molecular markers identify a deletion associated with quakingviable.* Genetics, 1992. **131**: p. 183-90.

15. King, T.R. and W.F. Dove, *Pleiotropic action of the murine quaking locus: structure of the qkv allele.* Mamm. Genome, 1991. **1**: p. 47-52.

16. Kondo, T., et al., *Genomic organization and expression analysis of the mouse qkI locus.* Mamm. Genome, 1999. **10**: p. 662-9.

17. Vernet, C. and K. Artzt, *STAR, a gene family involved in signal transduction and activation of RNA.* Trends Genet., 1997. **13**: p. 479-84.

18. Zorn, A.M. and P.A. Krieg, *The KH domain protein encoded by quaking functions as a dimmer and is essential for notochord development in Xenopus embryos.* Genes Dev., 1997. **11**: p. 2176-90.

19. Chen, T., et al., *Self-association of the single-KH-domain family members Sam68, GRP33, GLD-1, and Qk1: role of the KH domain.* Mol. Cell Biol., 1997. **17**: p. 5707-18.

20. Wu, J., et al., *The quaking I-5 protein (QKI-5) has a novel nuclear localization signal and shuttles between the nucleus and the cytoplasm.* J. Biol. Chem., 1999. **274**: p. 29202-10.

21. Pilotte, J., D. Larocque, and S. Richard, *Nuclear translocation controlled by alternatively spliced isoforms inactivates the QUAKING apoptotic inducer.* Genes Dev., 2001. **15**: p. 845-58.

22. Hardy, R.J., et al., *Neural cell type-specific expression of QKI proteins is altered in quakingviable mutant mice.* J. Neurosci., 1996. **16**: p. 7941-9.

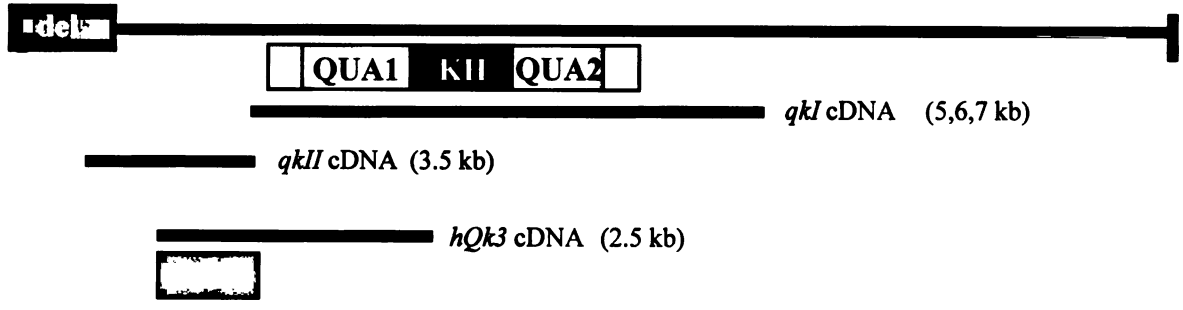
23. Saccomanno, L., et al., *The STAR protein QKI-6 is a translational repressor.* Proc. Natl. Acad. Sci. U S A, 1999. **96**: p. 12605-10.

24. Wang, L.L., S. Richard, and A.S. Shaw, *P62 association with RNA is regulated by tyrosine phosphorylation.* J. Biol. Chem., 1995. **270**: p. 2010-3.

25. Barlat, I., et al., *A role for Sam68 in cell cycle progression antagonized by a spliced variant within the KH domain.* J. Biol. Chem., 1997. **272**: p. 3129-32.

26. Crittenden, S.L., et al., *A conserved RNA-binding protein controls germline stem cells in Caenorhabditis elegans.* Nature, 2002. **417**: p. 660-3.

A


del
QUA1
KII
QUA2
qkI cDNA (5,6,7 kb)
qkII cDNA (3.5 kb)
hQk3 cDNA (2.5 kb)

B

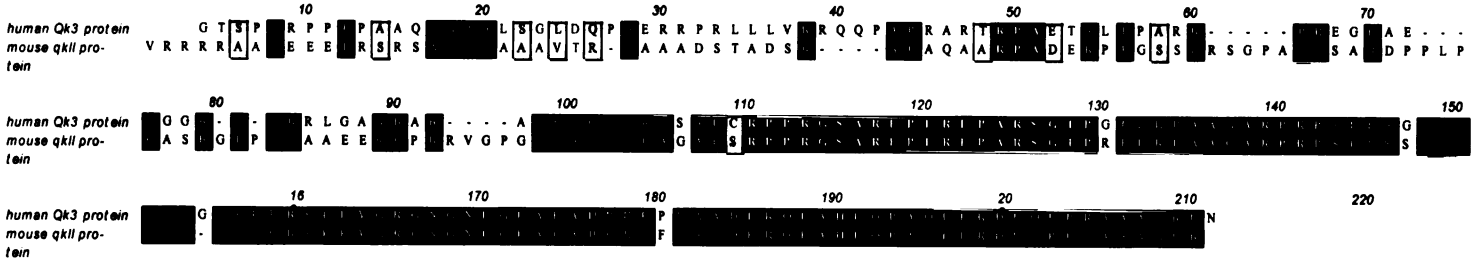
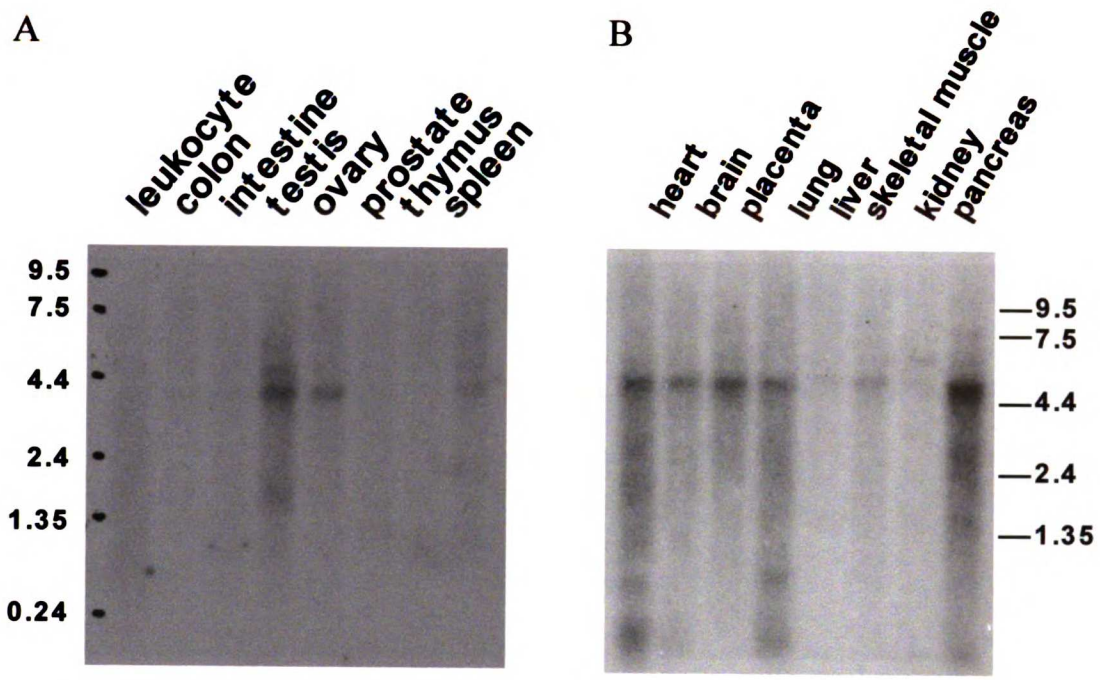

human Qk3 protein
mouse qkII protein
human Qk3 protein
mouse qkII protein
human Qk3 protein
mouse qkII protein

Figure 1


A
leukocyte
colon
intestine
testis
ovary
prostate
thymus
spleen
9.5
7.5
4.4
2.4
1.35
0.24
B
heart
brain
placenta
lung
liver
skeletal muscle
kidney
pancreas
9.5
7.5
4.4
2.4
1.35

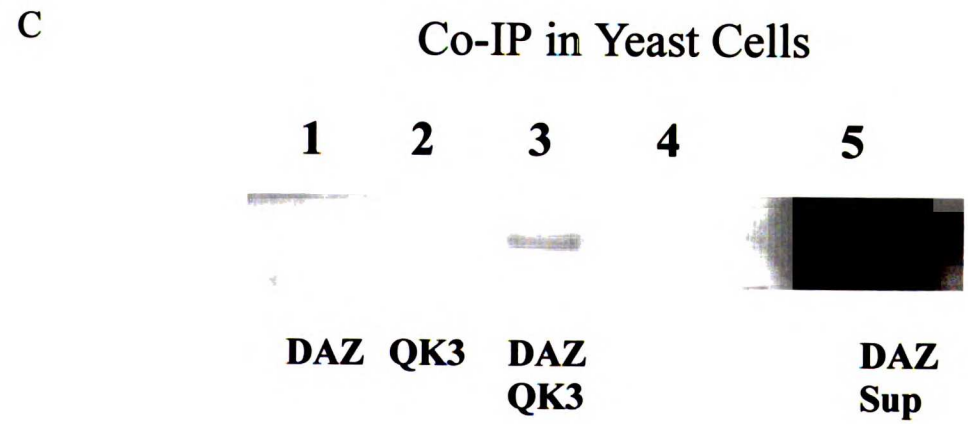

C
Co-IP in Yeast Cells
1
2
3
4
5
DAZ
QK3
DAZ
QK3
DAZ
Sup

Figure 2

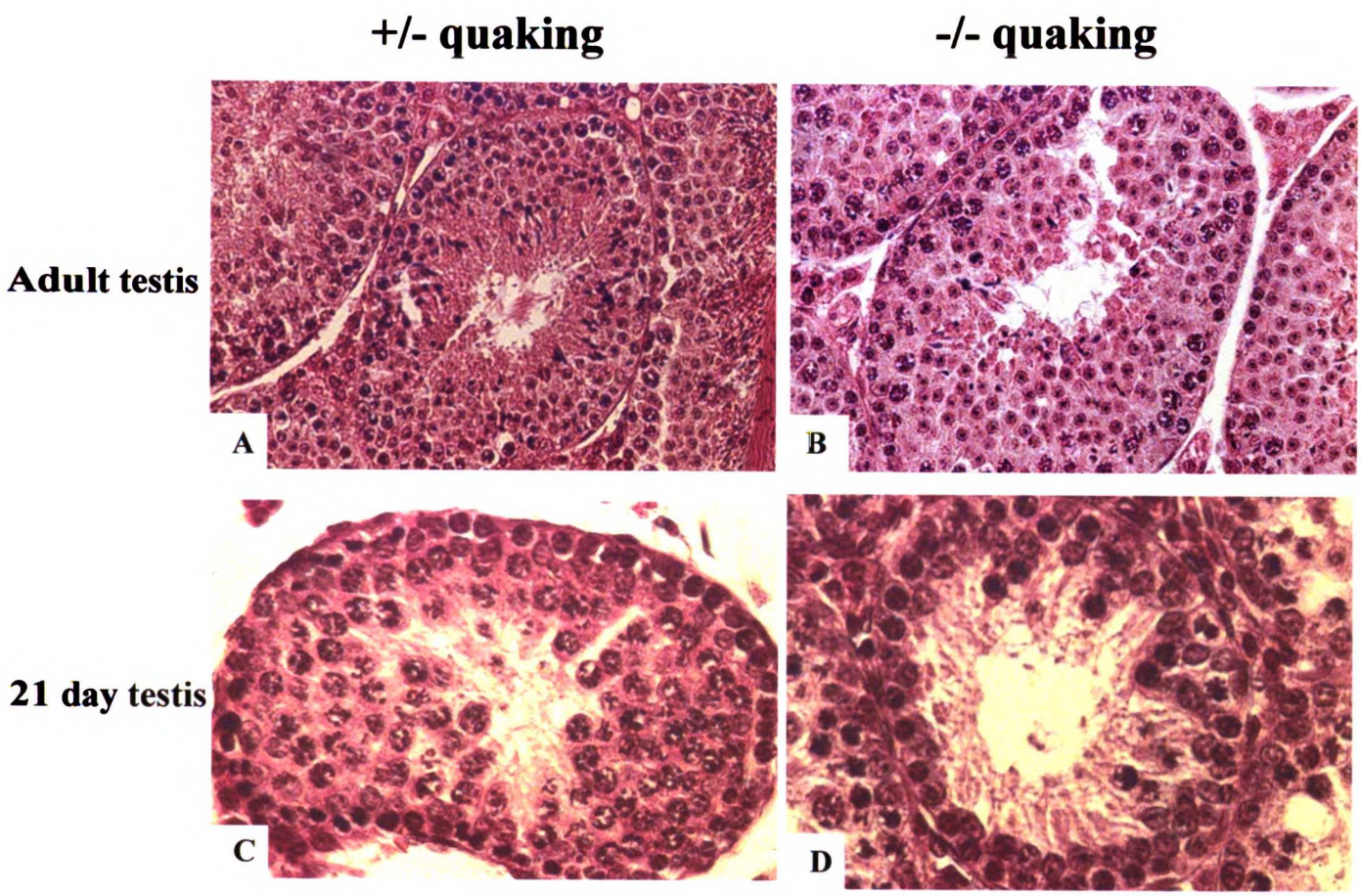


Figure 3

Model for Quaking Proteins Function in Testis

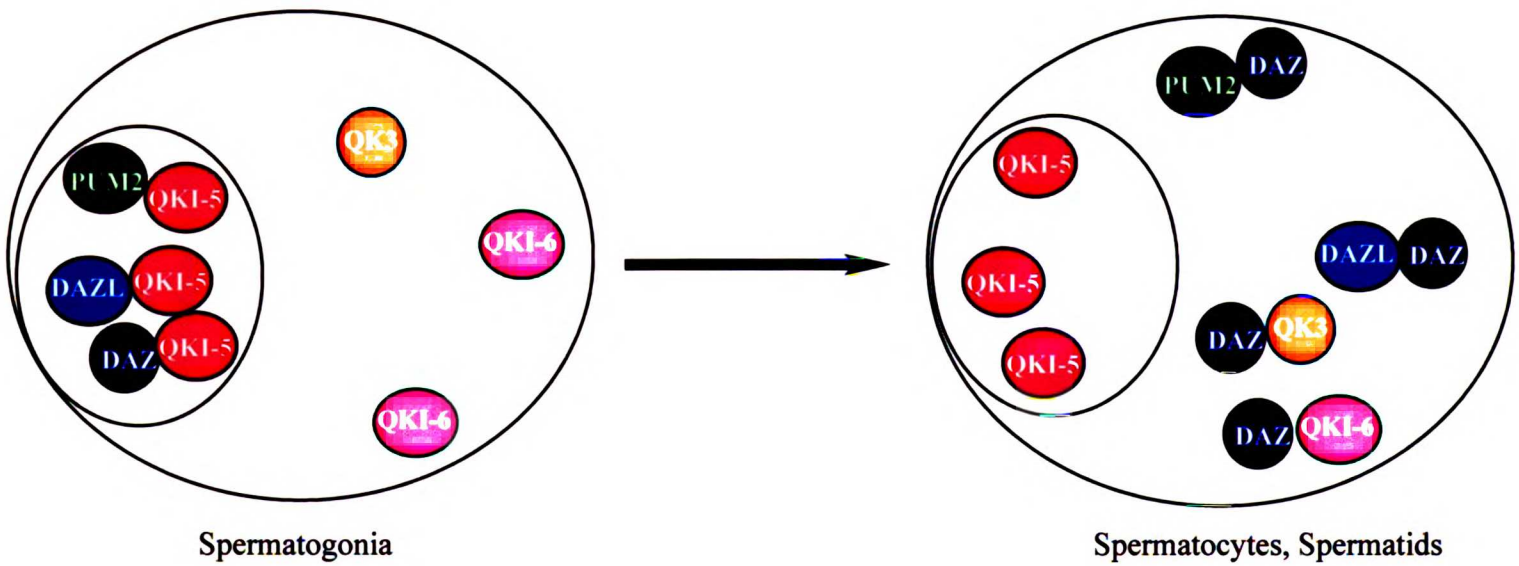


Figure 4

Chapter 6

Conclusion

Conclusion

Initial theme of DAZ-associated candidate proteins

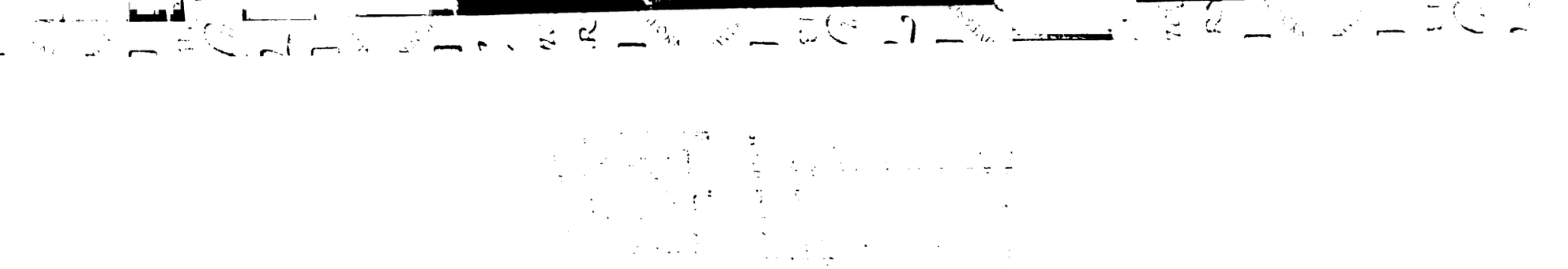
The identification of *DAZ*[1] and its homologue, *DAZL*[2, 3] was a significant step towards understanding the genetics of human germ cell development. However, to better understand the molecular function of the DAZ and DAZL proteins, the identification of RNA substrates and interacting cofactors is essential. We identified seven candidate genes that encode proteins that potentially interact with DAZ[4]. Initial information based on sequence analysis and data acquired in model organisms allowed us to suggest the following hypotheses about our candidate genes. First, five of the seven candidate genes encoded proteins that contained domains that could potentially bind RNA. Based on the abilities of DAZ and DAZL proteins to bind RNA[5, 6], these initial data allowed us to suggest that DAZ may interact with cofactors, which can additionally bind RNA for their functions in spermatogenesis. Secondly, four of the seven candidate genes were homologues of genes from a model organism that had been shown genetically to be associated with infertility[1, 7-10]. This information allowed us to hypothesize that genes which encode proteins that potentially interact with DAZ may be part of a conserved germ cell machinery. Moreover, this machinery may be required for normal germ cell development in humans. Further characterization of five of the seven candidate genes has allowed us to build a testable molecular model for the function of DAZ, DAZL, and DAZ-associated cofactors at different stages of germ cell development.

A model for the role of DAZ and associated cofactors at different stages in germ cell development

Evidence from this research, along with that from previous research in humans and other organisms, supports functional roles for DAZ and DAZ-associated proteins early in germ cell development and later at meiosis. We identified and characterized DAZ-associated cofactors (PUM2, DZIP1, DZIP2) that are present specifically in ES cells and germ cells and colocalize with DAZ and/or DAZL proteins in fetal and adult germ cells. In addition, we identified the DAZ-associated cofactor, BOL, which is expressed exclusively in the testis and colocalized with DAZ and DAZL proteins at the onset of meiosis. We integrated the information discovered in our research study with data acquired on the *Pumilio* and *Boule* genes from model organisms to build a molecular model for DAZ and DAZL function in human germ cells. This model will serve to direct future experiments on germ cell development in men and women. The essence of this model is shown in Figure 1 and is based on observations outlined in previous chapters. Essentially, DAZ, and DAZ-associated cofactors function in a combinatorial fashion to regulate translation of transcripts required at specific stages in germ cell development; moreover, the disruption of these associated cofactors could lead to a depletion of the germ cell population and/or arrest of meiosis, depending on its stage-specific function.

Future Directions

This is an exciting time to study germ cell development in men and women. We can now begin to understand, at the molecular level, the role of potential fertility factors, such as the members of the *DAZ* gene family and their associated cofactors in germ cell



development. The research described here has increased our understanding of protein interactions in germ cell development and yet many questions remain to be addressed: Potential RNA substrates that DAZ may regulate have been identified. However, the lack of *in vivo* assays, which address the loss of function of DAZ and DAZ-associated cofactors, are critical for further dissection of the human germ cell pathways. I suggest that RNAi in cultured mammalian ES cells and spermatogonial stem cells holds promise as a tool for phenotypic analysis for DAZ and DAZ-associated cofactors *in vivo*. By analyzing the mRNA expression of disrupted DAZ and DAZ-associated cofactors (PUM2, DZIP1, DZIP2, BOL, etc), utilizing cDNA microarrays, it is possible to identify direct and indirect downstream targets regulated by DAZ. In addition, *in vitro* assays (mobility shift, etc) will help to determine if DAZ and DAZ-associated cofactors can bind as a cluster on RNA and will force the development of innovative assays that address the role of these interactions *in vivo*.

I am personally excited about the new connections that have been made between the genes required for germ cell development in model organisms and in humans. The role of genes that have been conserved through evolution and genes that have a more recent evolutionary history will be elucidated in the years to come. It is likely that the results of studies such as these will benefit the significant portion of the population that suffers from reproductive health problems.

Figures

Figure 1 Model of DAZ and DAZ-associated proteins in ES cells and germ cells. Some DAZ-associated proteins are expressed early in embryonic stem cells and through germ cell development. In contrast, additional DAZ-associated proteins are expressed at the onset of meiosis in germ cells. This model suggest that DAZ and DAZ-associated proteins interact in a combinatorial fashion at different stages in germ cell development to regulate translation of transcripts required for pluripotency/totipotency and normal oogenesis and spermatogenesis.

References

1. Reijo, R., et al., *Diverse spermatogenic defects in humans caused by Y chromosome deletions encompassing a novel RNA-binding protein gene*. Nat. Genet., 1995. **10**: p. 383-93.

2. Saxena, R., et al., *The DAZ gene cluster on the human Y chromosome arose from an autosomal gene that was transposed, repeatedly amplified and pruned*. Nat. Genet., 1996. **14**: p. 292-9.

3. Yen, P.H., N.N. Chai, and E.C. Salido, *The human autosomal gene DAZLA: testis specificity and a candidate for male infertility*. Hum. Mol. Genet., 1996. **5**: p. 2013-7.

4. Moore, F.L., Jaruzelska, J., Fox, M.S., Urano, J., Firpo, M.T., Turek, P.J., Dorfman, D.M., Reijo-Pera, R.A., *A Human Homolog of Drosophila Pumilio is Expressed in Embryonic Stem Cells and Germ Cells and Interacts with DAZ (Deleted in AZoospermia) and DAZ-Like Proteins*. Submitted to PNAS, 2002.

5. Jiao, X., P. Trifillis, and M. Kiledjian, *Identification of target messenger RNA substrates for the murine deleted in azoospermia-like RNA-binding protein*. Biol. Reprod., 2002. **66**: p. 475-85.

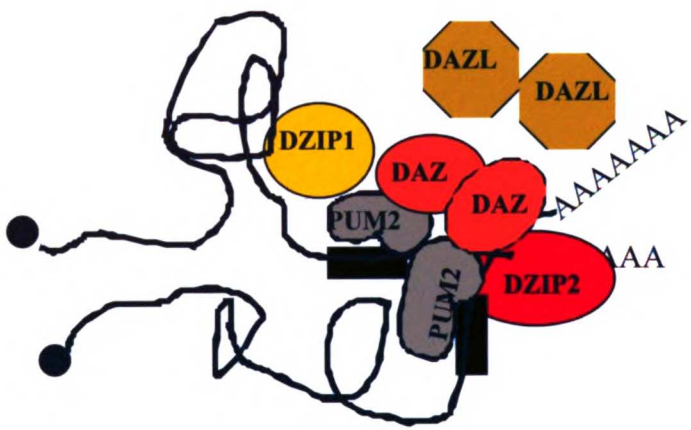
6. Venables, J.P., M. Ruggiu, and H.J. Cooke, *The RNA-binding specificity of the mouse Dazl protein*. Nucleic Acids Res., 2001. **29**: p. 2479-83.

7. Ruggiu, M., et al., *The mouse Dazla gene encodes a cytoplasmic protein essential for gametogenesis*. Nature, 1997. **389**: p. 73-7.

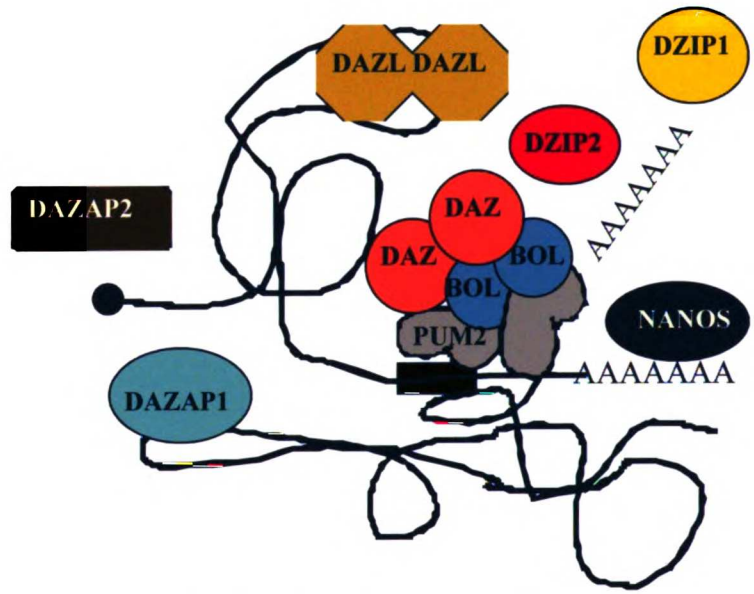
8. Forbes, A. and R. Lehmann, *Nanos and Pumilio have critical roles in the development and function of Drosophila germline stem cells*. Devel., 1998. **125**: p. 679-690.

9. Chubb, C., *Oligotriche and quaking gene mutations. Phenotypic effects on mouse spermatogenesis and testicular steroidogenesis*. J. Androl., 1992. **13**: p. 312-7.

10. Ebersole, T.A., et al., *The quaking gene product necessary in embryogenesis and myelination combines features of RNA binding and signal transduction proteins*. Nat. Genet., 1996. **12**: p. 260-5.

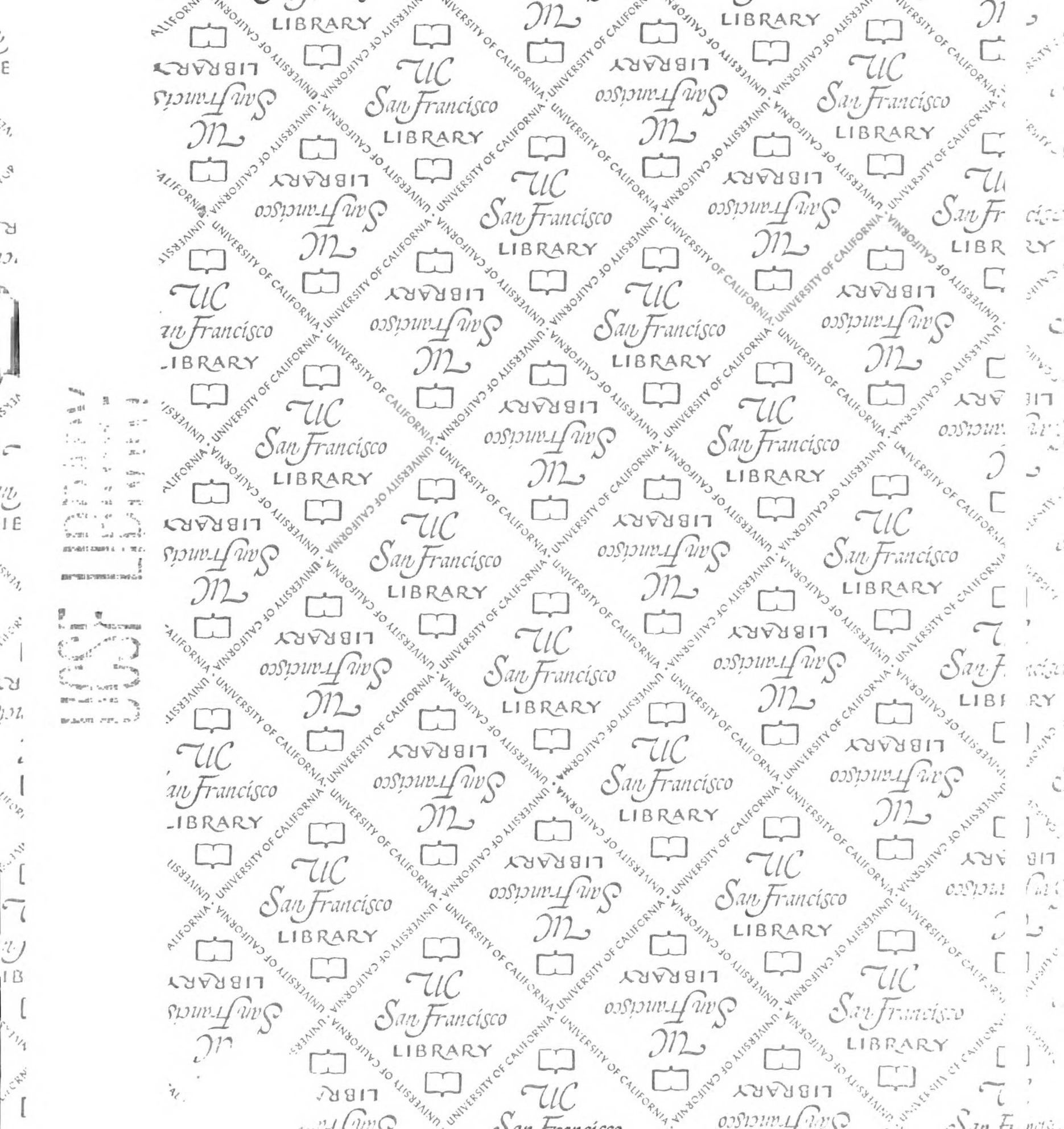


Stem Cell & Germ Cell Development



Progression through meiosis

Figure 1


UC
San Francisco
LIBRARY
UNIVERSITY OF CALIFORNIA · UNIVERSITY OF CALIFORNIA