Haploinsufficiency of the germ cell-specific nuclear RNA binding protein hnRNP G-T prevents functional spermatogenesis in the mouse

Ingrid Ehrmann1, Caroline Dalgliesh1, Aikaterini Tsaousi1, Maria Paola Paronetto2,3, Bettina Heinrich6, Ralf Kist1, Paul Cairns1, Weiping Li1, Christian Mueller1,8, Michael Jackson1, Heiko Peters1, Karim Nayernia1,8, Philippa Saunders4, Michael Mitchell5, Stefan Stamm7, Claudio Sette2,3 and David J. Elliott1,*

1Institute of Human Genetics, Newcastle University, Central Parkway, Newcastle NE1 3BZ, UK, 2Department of Public Health and Cell Biology, Section of Anatomy, University of Rome ‘Tor Vergata’, Rome, Italy, 3IRCSS Fondazione Santa Lucia, Rome, Italy, 4MRC Reproductive Sciences Unit, Centre for Reproductive Biology, 49 Little France Crescent, Edinburgh EH16 4SB, Scotland, UK, 5Génétique moléculaire de la spermatogenèse, Inserm U491, Faculté de médecine, 27, bd Jean Moulin 13385, Marseille, France, 6Institut fur Biochemie, Emil-Fischer-Zentrum, Friedrich-Alexander Universitat Erlangen-Nurnberg, Fahrstrasse 17 91054, Erlangen, Germany, 7Department of Molecular and Cellular Biochemistry, College of Medicine, University of Kentucky, 741 South Limestone, Lexington, KY 40536-0509, USA and 8North East Stem Cell Institute (NESCI), Newcastle, UK

Received May 1, 2008; Revised May 29, 2008; Accepted June 16, 2008

Human HNRNP GT, encoding the protein hnRNP G-T, is one of several autosomal retrogenes derived from RBMX. It has been suggested that HNRNP GT functionally replaces the sex-linked RBMX and RBMY genes during male meiosis. We show here that during normal mouse germ cell development, hnRNP G-T protein is strongly expressed during and after meiosis when proteins expressed from RbmX or RbmX-like genes are absent. Amongst these RbmX-like genes, DNA sequence analyses indicate that two other mouse autosomal RbmX-derived retrogenes have evolved recently in rodents and one already shows signs of degenerating into a non-expressed pseudogene. In contrast, orthologues of Hnrnpgt are present in all four major groups of placental mammals. The sequence of Hnrnpgt is under considerable positive selection suggesting it performs an important germ cell function in eutherians. To test this, we inactivated Hnrnpgt in ES cells and studied its function during spermatogenesis in chimaeric mice. Although germ cells heterozygous for this targeted allele could produce sperm, they did not contribute to the next generation. Chimaeric mice with a high level of mutant germ cells were infertile with low sperm counts and a high frequency of degenerate seminiferous tubules and abnormal sperm. Chimaeras made from a 1:1 mix of targeted and wild-type ES cell clones transmitted wild-type germ cells only. Our data show that haploinsufficiency of Hnrnpgt results in abnormal sperm production in the mouse. Genetic defects resulting in reduced levels of HNRNP GT could, therefore, be a cause of male infertility in humans.

INTRODUCTION

The heterogeneous ribonucleoproteins (abbreviated hnRNPs) are a family of nuclear RNA binding proteins with roles in nuclear RNA metabolism including the regulation of pre-mRNA splicing (1). HnRNP G-T (an acronym of heterogeneous ribonucleoprotein G expressed in the Testis) is a testis-specific hnRNP encoded by an autosomal retrogene

*To whom correspondence should be addressed. Tel: +1 912418694; Fax: +1 912418666; Email: david.elliott@ncl.ac.uk

© The Author 2008. Published by Oxford University Press. All rights reserved.
For Permissions, please email: journals.permissions@oxfordjournals.org
HNRNPGT) on human chromosome 11p15. Sequence homology suggests that the HNRNPGT gene is derived from retrotransposition of a transcript from the X-linked RBMX gene (which encodes the nuclear hnRNP G protein) (2). X-linked retrogenes are thought to be important during pachytene of male meiosis when the sex chromosomes of mammals are inactivated by a process called Meiotic Sex Chromosome Inactivation and relocate in a nuclear structure called the XY body (3–5). The XY body is repressive for transcription, but a number of autosomal retrogenes derived from X-linked genes are meiotically expressed. These autosomal retrogenes may encode either direct functional replacements for essential X-linked housekeeping genes such as PGK1 which are turned off during meiosis, more specialized germ cell proteins with distinct functions from their X-linked counterparts, or possibly play a post-meiotic role to normalize gene expression between X- and Y-chromosome containing round spermatids (6).

Consistent with a role in meiosis, human hnRNP G-T protein is highly expressed in the nuclei of primary spermatocytes and in round spermatids (2). However in addition to HNRNPGT, there are nine other retrogene copies of RBMX in the human genome, including a full length copy on chromosome 1 which is more closely similar to RBMX (RBMXL1) than HNRNPGT, and a slightly truncated version on chromosome 9 (RBMXL9) (7). Expression analysis of these genes using RT–PCR has shown that RBMXL1 is expressed ubiquitously (including within the testis) while RBMXL9 expression is restricted to brain and testis. The presence of two other full-length and more closely related RBMX retrogenes in the human genome (7) and at least one in the mouse (8) suggest the hypothesis that these other retrogenes might in fact replace RBMX gene function during meiosis, and that Hnrgpt might have evolved distinct functions.

Because of its testis-specific expression, defects in the HNRNPGT gene might play a role in human infertility. Sequence variations in the HNRNPGT gene have been identified in men with impaired fertility, and hnRNP GT protein is not expressed in men showing incomplete spermatogenesis (9,10). HNRNPGT is also implicated with a role in infertility through its sequence homology to the Y-linked RBMY genes (2). RBMY genes were originally identified by positional cloning as candidates for a Y chromosome gene deleted in some infertile men called the Azoospermia factor (AZF) and microdeletions of the AZFb region of the human Y chromosome prevent expression of RBMY protein (11,12). Deletions which reduce mouse Rbmy gene copy number are also associated with sperm abnormalities (13) and Rbmy genes are conserved in all mammals suggesting a fundamental role in male germ cell development which was established early in mammalian evolution (14). However, because mouse Rbmy is a multiple copy gene on the Y chromosome, it is technically difficult to inactivate by gene targeting in mice.

Here we have analysed the genetic history and expression patterns of Hnrgpt compared with Rbmx and other Rbmx-derived retrogenes, and tested the function of Hnrgpt in mouse germ cell development by gene targeting in mice. Our results point to a unique and important function for hnRNP G-T protein in meiotic and post-meiotic germ cell development.

RESULTS

HNRNPGT orthologues are present in all superorders of placental mammals and are under positive selection

An antiserum raised against human hnRNP G-T protein identified a testis-specific protein in mouse and bull suggesting the existence of an hnRNP G-T orthologue in these species (2). If the HNRNPGT gene indeed performs a basic biological function during germ cell development, this function would be expected to be conserved between species. To test this, we searched for genes in other species which are both retrogenes (i.e. do not contain introns) and map to genomic regions syntenic to the human HNRNPGT gene (adjacent to the NLRP14 gene within the genomic sequence). Using these criteria, we found putative HNRNPGT orthologues in species representing three of the four super-orders of placental mammals (15); Laurasiatheria (mouse and rat), Eutheria (elephant and tenrec). A retrogene was also identified in one Xenarthran species (armadillo), although because of the incomplete genome sequence a syntenic relationship with the human HNRNPGT gene could not be confirmed. This species distribution is consistent with HNRNPGT being created by a single retrotransposition event before the radiation of placental mammals. Using the same criteria, we were unable to identify HNRNPGT orthologues in any marsupial genome.

To investigate the relationship between putative HNRNPGT orthologues and RBMX further we used the sequences identified, together with RBMX sequences from the same species (where available) to construct phylogenetic trees with the RBMX orthologue from chicken serving as an outgroup (Fig. 1). In this tree, the HNRNPGT and RBMX sequences form two distinct clades with 100 and 97% bootstrap support, respectively. In both clades, some branching orders are inconsistent with known phylogenetic relationships (15), most noticeably the positioning of the tenrec and cow HNRNPGT genes. The position of these genes within the tree does not have high bootstrap support, highlighting the uncertainty of the inferred relationships. What is clear is that the basic topology of the tree is consistent with the HNRNPGT sequences being created from a single retrotransposition event which occurred before the radiation of eutherian mammals. Importantly, the branch lengths within the HNRNPGT and RBMX clades are different. This could be explained by a strong selection for conservation of function acting upon RBMX, with HNRNPGT evolving faster. More rapid evolutionary change might indicate HNRNPGT is under less selective pressure, or might be because of the increased evolutionary rate of genes involved in reproductive processes (16). The branch lengths in the phylogenetic tree show that HNRNPGT is less conserved than RBMX. This is evident from the rate of non-synonymous substitutions (Ka) between human and mouse RBMX (0.006) and human and mouse HNRNPGT (0.109) and also reflected in the Ka/Ks ratio which is 10 times higher for HNRNPGT (0.20) than RBMX (0.02) (Supplementary Material, Tables S1a and 1b). The observed Ka/Ks ratio of less than 1 demonstrates that the coding sequence of HNRNPGT has been conserved, but indicates that, compared to RBMX, it has been subject to reduced selection pressure and/or positive selection for amino acid substitutions (see legend to Supplementary Material, Table S1).
The mouse genome contains two additional Rbmx-derived retrogenes which have evolved in the rodent lineage but only one of them is expressed

Our gene search in the mouse genome revealed two additional full-length mouse Rbmx-derived retrogenes (Chromosome 14 copy, Rbmxrt14: ENSMUSG00000049235; Chromosome 8 copy, Rbmxrt8: ENSMUSG00000037070) which could potentially encode proteins which are greater than 95% identical to RBMX (Fig. 2: by comparison Hnrnpgt encodes a protein which is only around 75% identical to Rbmx). Neither Rbmxrt8 nor Rbmxrt14 are conserved in human, but an orthologue of Rbmxrt8 is found on rat chromosome 19 (Ensembl gene identification number ENSRNOG00000012138). This means that Rbmxrt8 and Rbmxrt14 evolved in the rodent lineage respectively before and after the mouse/rat divergence. Comparative analysis shows that Rbmxrt14 is a retroposed pseudogene (legend to Fig. 1) derived from a spliced Rbmxrt8 transcript bearing the shorter form of the 5'-UTR (see legend to Supplementary Material, Fig. S1). Rbmxrt8 is under greater selective constraint than Rbmxrt14 (Supplementary Material, Table S2a and b) and there are many Rbmxrt8 ESTs in the databases from a range of tissues, although we were unable to find a single EST derived from Rbmxrt14. Taken together, these data suggest that Rbmxrt8 is expressed and Rbmxrt14 is not.

Mouse hnRNP G-T is exclusively expressed after the initiation of meiosis

The above analysis identified an HNRNPGT gene (ENSMUSG00000073894, also known as Rbmxl2) located on a region of mouse chromosome 7 which is syntenic to human chromosome 11p15 where human HNRNPGT is located. Consistent with ENSMUSG00000073894 being the mouse orthologue of human HNRNPGT, Northern blotting identified a single transcript of ~1500 nucleotides from ENSMUSG00000073894 exclusively in mouse testis (Fig. 3 and data not shown). In addition, we raised specific antisera to mouse hnRNP G-T and demonstrated by western blotting that the protein is also testis-specific (Supplementary Material, Fig. S2). Northern blotting of RNA isolated from different mouse tissues detected expression of Rbmx, Rbmxrt8 and Hnrnpgt within the testis, along with expression of the testis-specific Rbmy gene. While Hnrnpgt expression was testis-specific, co-expression of the Rbmx and Rbmxrt8 genes was particularly evident in thymus, spleen, testis and brain which are tissues containing extensive alternative mRNA splicing programmes (Fig. 3). Two splice variants of the Rbmx transcript with alternative 3' ends (8) were also differentially expressed between tissues. Of these Rbmx alternative transcripts, splice variant 2 was highly expressed in the testis but splice variant 1 was not detectably expressed, although it has been detected by RT–PCR and was originally isolated from a testis cDNA library (8).

Since the Rbmxrt8 protein sequence is more similar to RBMX than the hnRNP G-T protein sequence and the Rbmx, Rbmxrt and Hnrnpgt genes are co-expressed within the testis (Figs 2 and 3), we considered the possibility that RBMXRT8 protein might in fact provide a more similar replacement for RBMX protein function during meiosis. To test if the cognate proteins from these genes were co-expressed in the same cells during germ cell development, we carried out...
further experiments using our antisera to hnRNP G-T and RBMX. To differentiate between cell types in the testis, we used three approaches. First, on western blots containing protein extracted from purified germ cell populations isolated by elutriation (17), no hnRNP G-T protein expression was detected in spermatogonia but strong hnRNP G-T protein expression was detected in primary and secondary spermocytes, and somewhat weaker expression in round spermatids.

Figure 2. A family of RBMX proteins are encoded by the mouse genome. (A) The protein sequence of human (Hs) hnRNPG-T protein aligned with the conceptual protein sequences encoded by the open reading frames of mouse (Mm) Hnrgnpt, encoded by chromosome 7, Rbmx (hnRNPG, encoded by the X chromosome), Rbmxrt8 and Rbmxrt14 (encoded by chromosome 14). Conserved sequences within the N terminal RNA recognition motif which is found within each of these proteins are boxed and striped (RNP1), or stippled (RNP2). The 18 amino acid peptide which was used to generate and affinity purify the mouse HnRNP G-T specific polyclonal antibody is lightly shaded, and the actual peptide sequence is in bold. The peptide to which the antisera specific to RBMX, RBMXRT8 and RBMXRT14 proteins was raised is indicated by a black box. (B) Table showing percentage sequence identity (I) and similarity (S) between human and mouse hnRNP-GT proteins, and between different members of the mouse RBM protein family measured using a pairwise alignment algorithm (http://www.ebi.ac.uk/emboss/align/index.html).
In contrast, high RBMX and RBMXRT8 protein expression was detected in purified spermatogonia, but not in the later stages of spermatogenesis (Fig. 4A, upper panel). An Erk2 loading control confirmed equal levels of protein loading in each lane.

Secondly, we probed western blots containing testis protein from mice going through the first wave of germ cell development, which contain germ cells enriched in specific stages of spermatogenesis. HnRNP G-T protein expression was low until day 13 (when germ cells are initiating meiosis I), and thereafter strong (Fig. 4B, middle panel). Parallel western blots showed that RBMX/RBMXRT8 protein expression was high in the testes of 4-day-old mice (when spermatogonia represent the sole germ cells in the testis) and low later on (as spermatogonia become diluted by later stages of germ cell differentiation) (Fig. 4B, top panel).

Thirdly, we used the affinity purified α-mouse hnRNP G-T antiserum to examine the expression pattern of hnRNP G-T protein on mouse testis sections using indirect immunofluorescence. High levels of expression were detected in the nuclei of primary spermatocytes, and to a lesser extent in round spermatids (abbreviated Sp and Rtd, respectively, in Fig. 4C). There was no detectable expression in spermatogonia (abbreviated Spg: only DAPI fluorescent counterstain was visible, pseudocoloured blue in Fig. 4C). This pattern of expression identified by immunostaining was similar to that previously observed for human hnRNP G-T (2). RT–PCR experiments have reported expression of the human RBMX-related retrogenes RBMXL1 and RBMXL9 in the testis (7). The peptide used to raise the antisera against mouse RBMX protein is also encoded within the Open Reading Frames of human RBMX and the RBMXL1 and RBMXL9 retrogenes. Consistent with the mouse data above, immunohistochemistry on human testis sections showed strong expression of RBMX/RBMXL1/RBMXL9 protein only in pre-meiotic stages of human germ cell development and within the nuclei of somatic Sertoli cells (abbreviated as SC in Fig. 4D: in this section the brown corresponds to specific staining by the antibody while the blue is the haematoxylin counterstain).

These results indicate that the mouse hnRNPGT protein has a similar pattern of expression to its human orthologue (2) suggesting similar meiotic and post-meiotic functions for hnRNPGT protein in human and mouse. Hence, Hnrpngt has acquired a promoter which is active during and after meiosis. In contrast, the RBMX and RBMXRT8 proteins are only detectably expressed prior to meiosis. Because of its lack of expression in meiosis, RBmxrt8 would be unable to provide an alternative meiotic replacement function for Rbmx.

Mouse chimaeras made from TBV-2 ES cells containing a targeted disruption of the Hnrpngt gene did not transmit the targeted allele

Our expression analyses indicated that hnRNP G-T is the only source of an RBMX-like protein during and after meiosis in the mouse and human testis. To test if Hnrpngt gene expression is important for normal germ cell development, we inactivated one Hnrpngt allele by gene targeting in TBV-2 ES cells using a construct in which the majority of the Hnrpngt gene is replaced by the lacZ gene (Fig. 5C). A correctly targeted ES cell clone (2.1D3) was isolated and injected into blastocysts resulting in the generation of seven chimaeras (six male and one female). This distortion in the sex ratio of chimaerics is indicative of good quality chimaeras, as were high levels of coat chimaerism (18,19). However, over the course of 9 months of breeding, the five fertile chimaeric males and one female produced 291 black pups but no agouti pups (indicative of transmission of the...
These experiments suggested that germ cells which contained a heterozygous disruption of the \textit{HnrnpGt} gene were unable to contribute to the germline. This might have been because of either (i) haploinsufficiency of the \textit{HnrnpGt} gene or (ii) some other problem with the line of ES cells used in the experiment. A published analysis of ES cells which had successfully gone germline showed that a level of euploidy of greater than 50% is a good prognostic indicator of germline transmission (18). We found that more than 90% of the cells in clone 2.1D3 were euploid (Table 1). The TBV-2 ES cell line we used in the above targeting experiments had also recently provided germline transmission of different targeted alleles in four other cases (20,21). However, in order to eliminate the possibility that the failure to transmit the targeted allele was the result of an undetected problem with our TBV-2 cell line, we carried out a further round of gene targeting in a second ES cell line E14.1. After re-targeting, we chose two new independently targeted ES cell clones 2.1G5 and 3.1B11 for further analysis and characterized these cell lines using Southern blot analysis to confirm precise integration by our targeting vector (Fig. 5A and B). The Southern blot also showed that clone 2.1G5 contains a homogeneous population of ES cells containing the targeted allele (Fig. 5A and B). Eighty-seven percent of cells within the 2.1G5 clone had a normal euploid karyotype (Table 1). Clone 2.1G5 was used to generate chimaeras which were again good quality according to the above criteria. Three high percentage coat colour chimaeras C1, C5 and C8 derived from the 2.1G5 cell line were fertile (Supplementary Material, Table S4). We then assessed the level of chimaerism at the DNA level using a strain-specific PCR (see methods) and detected a contribution of 10–15% of 129 genotype cells within the testes of chimaeras C1 and C8 (Fig. 6).
Despite this level of chimaerism for 129-derived cells, C1 and C8 produced between them in excess of 200 Bl6 pups and only 7 agouti pups. Although the Mendelian expectation is that the transgene should be transmitted to 50% of the offspring, none of the pups contained the targeted allele when assayed by either genomic PCR or Southern blotting (not shown).

To establish if 129 cells containing the targeted disruption of HnrnpGt were able to contribute efficiently to the germline, we carried out immunohistochemistry with a rabbit polyclonal antibody specific to the bacterial LacZ protein. Strong LacZ protein expression was detected within 10–15% of seminiferous tubules from chimaeras C1 and C8 (Fig. 6). When LacZ was expressed, it recapitulated expression detected for the endogenous hnRNP G-T protein: no LacZ staining was detected within spermatogonia; however, strong expression was detected in meiotic and post-meiotic germ cells (compare Fig. 7D with Fig. 4C). No immunohistochemical signal was detected in the testis of a wild-type C57Bl6 mouse (Fig. 7C), or when normal rabbit serum was used instead of the primary anti-LacZ antibody (not shown). Chimaera C5 was particularly informative. Although C5 contained a 30% ES129 contribution assayed by the strain-specific PCR (Figs 6 and 8C) or LacZ staining (Figs 6 and 7D), all 79 pups sired by this chimaera had black coat colour. To establish why the ES cell-derived sperm produced from C5 did not sire offspring, we compared the relative proportions of each genotype in genomic DNA from the testis, sperm from the epididymus of the chimaera and sperm collected from the uterus of a plugged DBA strain female. ES129-derived cells were detectable in the testes of C5 (Fig. 8C, lanes 5 and 6), at slightly reduced levels relative to C57Bl6 sperm in the epididymus (Fig. 8C, lanes 3 and 4) and then were undetectable as sperm within the uterus (Fig. 8C, lane 2: notice in this lane that the PCR product from the C57Bl6 sperm is still clearly visible). However, although we could not detect 129-derived sperm by strain-specific PCR, the targeted allele could be detected by gene-specific PCR (see Methods) in uterine sperm from chimaera C5 suggesting the presence of sperm cells containing the targeted allele at low levels (Fig. 8G, lane 7).

Table 1. Karyotype analysis of each of the ES cells used in this study

<table>
<thead>
<tr>
<th>ES cell clone</th>
<th>ES cell line</th>
<th>Euploid karyotypes (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1D3</td>
<td>TBV-E2</td>
<td>93</td>
</tr>
<tr>
<td>2.1G5</td>
<td>E14</td>
<td>87</td>
</tr>
<tr>
<td>3.1B11</td>
<td>E14</td>
<td>70</td>
</tr>
</tbody>
</table>

Figure 5. Targeting of the HnrnpGt gene in ES cells by homologous recombination. (A) Southern blot of genomic DNA from E14 ES cells, and the 3.1B11 and 2.1G5 clones after digestion with BgIII. The blot was probed with the 5′ probe and developed by autoradiography. (B) Southern blot of genomic DNA from E14 ES cells, and the 3.1B11 and 2.1G5 clones after digestion with BsrGI. The blot was probed with the 3′ probe and developed by autoradiography. (C) Map of the wild-type HnrnpGt genomic locus (above) and the targeted loci (below) showing the positions of the HnrnpGt gene and LacZ transgene, restriction sites used in the analysis and location of the 5′ and 3′ probes. Above each restriction map are shown the fragments detected by the 3′ probe in the wild-type and targeted locus, and below the restriction map are shown the fragments detected by the 5′ probe. Southern blots of restriction digested genomic DNA from clone 3.1B11 confirmed the genomic structure of the wild-type and targeted alleles within the ES cells. However, these blotting experiments detected the restriction fragment corresponding to the wild-type allele slightly more strongly than that of the targeted allele. Since there is one targeted allele and one wild-type allele per ES cell, the theoretically expected ratio for a pure clone of heterozygote ES cells containing a targeted allele should be 50:50. Consistent with this, phosphorimager analysis confirmed that clone 2.1G5 gave a ratio of 50:50. Clone 3.1B11 gave a ratio of 70:30. This is very close to the expected ratio of 75:25 if clone 3.1B11 contained one wild-type (untargeted) ES cell for every targeted ES cell (in this case, there would be three wild-type alleles for every targeted allele). This indicated that clone 3.1B11 contains a mixture of targeted ES cells and untargeted ES cells at an approximate ratio of 1:1.
A group of infertile high percentage chimaeric mice made from targeted E14.1 ES cells produced abnormal sperm

Another group of 2.1G5-derived chimaeric males were infertile (C3, C4 and C6) but produced sperm. Sections from the testes of these mice showed a mixed phenotype: some seminiferous tubules were sertoli cell only (SCO), but others contained a complete complement of germ cells (Fig. 7B and E). The level of testicular chimaerism detected within chimaeras C3, C4 and C6 was quite high (between 27 and 42% 129 genotype using the strain-specific genomic PCR) and in each case was comparable to that detected in the kidney (Fig. 8A and B). The majority of seminiferous tubules from chimaeras C3, C4 and C6 which contained germ cells also stained positive for LacZ protein, indicating that the germ cells contained the targeted gene disruption of Hnmpt (Fig. 7B). In some cases, residual LacZ positive germ cells were also visible in otherwise SCO tubules, suggesting these SCO tubules had originally contained germ cells expressing the LacZ gene (indicated by an arrowhead in Fig. 7E). Other tubules from C3, C4 and C6 which were devoid of germ cells also frequently contained groups of sloughed off cells which might be germ cell-derived (arrowed in Fig. 7E).

Although they were infertile, C3, C4 and C6 produced 129 ES cell-derived epididymal sperm which could be detected at a high level by strain-specific PCR amplification and agarose gel electrophoresis and plotted as a percentage. The percentage of LacZ positive tubules and SCO tubules are the means of between two and seven separate sections from different parts of the testes.

Some high percentage infertile mouse chimaeras derived from targeted E14.1 ES cells were azoospermic

Two further chimaeras (C2 and C7) were also infertile, had particularly small testes (the testes of chimaera 7 next to a C57Bl6 littermate is shown in Fig. 7A) and did not produce any sperm. Sections of these testes showed they contained seminiferous tubules apparently devoid of germ cells (Fig. 7F). Both chimaeras C2 and C7 had a 30–50% ES129 cell contribution to the testis relative to C57Bl6 at the level of genomic DNA (Fig. 6, Fig. 8A, lanes 3–4 and Fig. 8B, lanes 5–6). Not even C57Bl6 sperm were produced by chimaeras C2 and C7: no sperm were detected within the epididymus and only a DBA strain genomic PCR signal was detected in a uterine sample of female DBA mice plugged
by these chimaeras (Fig. 8A, lanes 2 and Fig. 8B, lane 4). We also tested genomic DNA extracted from the kidney to which the targeted ES129 cells had also contributed (Fig. 8A, lane 5 and Fig. 8B, lane 7).

Overall, these data suggested that despite a high contribution of targeted ES129 germ cells within the testes and epididymal sperm of chimaeras derived from 2.1G5, there was a failure to transmit the targeted allele, high levels of testicular

Figure 7. Testicular phenotype in chimaeric mice made from ES cell clone 2.1G5. (A) The two testes of chimaera 7 next to the testis of a C57B6 littermate (Scale bar = 2 mm). (B) A low resolution light micrograph of a section from chimaera C3 immunostained for LacZ (Scale bar = 20 μm). (C) High-resolution light micrograph of a testis section from a wild-type C57B6 littermate immunostained for LacZ. (D) High-resolution light micrograph of a testis section from chimaera C5 immunostained for LacZ. (E) High-resolution light micrograph of Chimaera C4 immunostained for LacZ. (F) High-resolution light micrograph of a testis section from Chimaera C7 immunostained for LacZ. Scale bars in (C)–(F) are equal to 50 μm.
Figure 8. Genetic contribution of tissues and sperm from chimeraic mice made from ES cells containing the targeted allele of Hnrnpgt. (A–F) Genomic DNA was isolated from tissue, epididymal sperm or uterine sperm from plugged DBA females, amplified by genomic PCR and analysed by agarose gel electrophoresis to distinguish the cellular contribution of each genotype in the starting sample. (G) Genomic DNA was analysed using PCR primers to detect the targeted locus.
degeneration, low sperm counts and abnormal sperm morphology (summarized in Fig. 6). Full breeding data for these chimaeras are given in Supplementary Material, Table S4.

Chimaeric mice made from a heterogeneous clone containing wild-type and targeted ES cells only transmitted the wild-type allele

The above results suggested that germ cells which were haploinsufficient for the *Hnrnpgt* gene were unable to undergo functional spermatogenesis. To test this hypothesis, we next characterized the second E14.1 derived ES cell clone 3.1B11. Southern blot analysis of ES cell clone 3.1B11 revealed different signal intensities for wild-type and targeted alleles indicating a mixture of targeted and untargeted ES cells at a ratio of 1:1 (Fig. 5A and B). Importantly, both the targeted and wild-type ES cells within this clone 3.1B11 were passaged and processed identically during our experiments so the untargeted ES cells provide an internal control for the effectiveness of the targeted ES cells to functionally contribute to the germline. Karyotype analysis indicated 70% of clone 3.1B11 cells were euploid (Table 1).

Six fertile chimaeras with a high level of chimaerism judged by coat colour (four males and two females) were obtained from clone 3.1B11 (Supplementary Material, Table S4). Of these, C11 had a relatively low proportion of LacZ positive tubules detected by immunohistochemistry: less than 20% of the genomic DNA in the testis was from the 129 strain of mouse and some of the sperm contained the transgene therefore, this mouse was very similar to C1 and C8 derived from clone 2.1G5 (Fig. 6 and Fig. 8D and G). C11 produced 135 C57Bl6 pups, but also four agouti pups which did not contain the targeted mutation. In contrast, two of the other high percentage coat colour chimaeras derived from clone 3.1B11 (C10 and C12) had a much higher testicular ES129 contribution but no LacZ positive seminiferous tubules (Figs 6 and 8E). This suggested that the testes of C10 and C12 were mainly populated by the non-targeted ES129 cells within clone 3.1B11. We confirmed this by gene-specific PCR which showed that the targeted allele was not detected within uterine sperm from either chimaera (Fig. 8G). The testis of chimaera C10 also contained some apoptotic tubules which might have originally contained cells with the targeted allele. Both C10 and C12 fathered almost exclusively agouti pups (254/258) (Fig. 6 and Supplementary Material, Table S4) which were consistently negative for the targeted allele (data not shown). Even though C15 had low levels of ES129-derived sperm, this mouse still produced 41 agouti pups (Fig. 6). Unlike the samples from C5 and C11, none of the sperm contained the targeting construct (Fig. 8G) and the level of 129 genotype sperm did not decrease in the uterus relative to the epididymus (Fig. 8F). These experiments performed with a mixed population of wild-type ES cells grown in parallel with targeted ES cells indicate that wild-type but not targeted ES cells could result in the generation of functional sperm in chimaeric mice.

**DISCUSSION**

Our expression analysis indicates that *Hnrnpgt* provides the only source of an hnRNPG-like protein throughout meiosis.
and post-meiotic germ cell development. This is a period of specialized cell division and intense differentiation which lasts up to 22 days in the mouse and considerably longer in humans. The expression of Hnrnpgt is thus likely to be fundamentally important for germ cell biology. Consistent with this the phylogenetic data in this paper indicate an ancient origin for the HNRNPGT gene prior to the divergence of the four main groups of placental mammals between 101 and 108 millions of years ago (15). To account for this evolutionary conservation, pairwise Ka/Ks analysis between human and mouse show that HNRNPGT genes are being maintained by positive selection. The conservation of HNRNPGT across evolutionary time is particularly compelling since our data also suggest that one of the other mouse RBMX genes (Rbmxrt14) is already starting to degenerate, despite its comparatively recent origin within the past million years in the rodent lineage. In addition to direct retrotransposition of Rbmx-derived transcripts generating autosomal retrogenes, our phylogenetic analyses indicate that Rbmxrt14 is derived directly from Rbmxrt8 rather than Rbmx itself. The evolution of a mouse retrogene derived from Rbmxrt8 indicates that this retrotransposition must have occurred in the germline, and most likely from an Rbmxrt8 transcript made within spermatogonia.

To test if the Hnrnpgt gene does play an important function in mouse spermatogenesis, we carried out gene targeting in ES cells and generated 22 chimaeras (19♂ and 3♀) from three independently targeted clones generated using two different starting lines of ES cells. Each of these chimaeras had a high level of chimaerism based on coat colour, normally a good prognostic indicator for germline transmission (19). The ES cells also had levels of euploidy much higher than that which would normally be consistent with germline transmission (18). Consistent with this, the lowest number of cells containing euploid chromosomes (70%) was actually observed in the cell line 3.1B11 which resulted in the highest number of agouti pups. Although we could detect 129 strain genomic contributions to the testis, germ cells and in some cases sperm, no agouti mice heterozygous for the targeted allele were born. In contrast, both E14.1-derived ES cell lines gave rise to wild-type agouti progeny, although at different ratios. These data suggest that the absence of one copy of the Hnrnpgt gene within targeted germ cells of chimaeric mice prevents transmission into the next generation. The phenotype of the chimaeras was variable, ranging from fertile (with low targeted ES129-contribution within the testis) to low sperm counts/abnormal sperm (with high targeted ES129-contribution within the testis) to azoospermic (with high targeted ES129-contribution within the testis). Chimaeric testes which contained a high percentage of heterozygosity for Hnrnpgt also had a high proportion of SCO tubules and were infertile. Some of the otherwise SCO tubules in these mice still contained a few residual germ cells which were LacZ-positive. Hence, the observed testicular degeneration in mice with high levels of chimaerism for Hnrnpgt targeted cells might have been progressive, and by extrapolation testes which were entirely devoid of germ cells might have initially contained germ cells heterozygous for the transgene. Consistent with this idea, one chimaera in our study (C4) had one small and one slightly larger testis. The progressive failure of spermatogenesis with age has been reported for a targeted mutation in another mouse gene (22) and is also implicated in some genetic cases of human male infertility (23). Although they did not transmit to the next generation, we could detect ES129-containing sperm in some of the 2.1G5-derived chimaeras. Noticeably, chimaera C5 had a much higher ES129 contribution relative to C57Bl6 than either chimaeras C11 and C15, but did not father any agouti pups. The experiments which showed that the amount of 129 cells in the uterine sperm decreased markedly in C5 relative to the epididymal sperm, whereas the wild-type 129 sperm which appeared in the uterus from C15 looked similar to the levels in the epididymus suggested a failure of the transgenic sperm to either survive in the uterus or leave the epididymus.

Overall, our data suggest that all the major stages of spermatogenesis can proceed within seminiferous tubules populated by germ cells containing a heterozygous deletion of Hnrnpgt, but functional sperm containing this deletion are not produced and there is a failure of seminiferous tubule survival in adult life. Hence deletion of Hnrnpgt is likely to be affecting the maintenance of germ cell differentiation in the adult chimaeras. Sperm produced in the infertile chimaeras were very abnormal in morphology, particularly in the sperm head. Sperm head abnormalities have also been reported for mice containing the Yd1 deletion which reduces Rbmy expression 50-fold (13). Decreases of Rbmy copy number might have a less severe phenotype than heterozygosity for Hnrnpgt possibly since reductions in Rbmy expression could be compensated for by active Rbmx and Rbmxrt8 expression in spermatogonia (13,24–26). For the targeted allele in our experiments, most of the coding region of Hnrnpgt is replaced by lacZ. This is hence very likely to be a null allele resulting in reduced hnRNP G-T expression in targeted germ cells.

Two other retrogenes (Jsd and tauCstF-64) derived from X-linked genes have been shown to be essential for mouse spermatogenesis (27–29). In both these cases, heterozygous mice were fertile. However, some other chimaeric mice harbouring heterozygous disruptions of specific genes encoding important germ cell proteins are infertile. These include chimaeric mice containing germ cells heterozygous for the Prm1 and Prm2 gene which encodes protamines responsible for packaging DNA in the sperm, in the Klh10 gene which is also expressed in spermiogenesis, and in the Pf20 and Cyp17 genes (30–33). Biallelic expression might also be important for other RNA binding proteins expressed in germ cell development. Heterozygous deletions within the gene encoding the RNA binding protein Dazl also led to reduced testis size (34). The DAZL protein is thought to be primarily a translational regulator in the cytoplasm of meiotic and post-meiotic cells in adult life, but is nuclear in prenatal life (35). Why might normal bi-allelic expression of hnRNP GT protein be essential for spermatogenesis? Unlike mitosis, transcription is very prolific in meiosis, which will generate an abundant source of pre-mRNA for splicing (36) and hnRNP and hnRNP GT proteins are thought to function as a splicing factors (37,38). Splicing is thought to be controlled in part by the nuclear concentration of RNA binding proteins (39–41). Reductions in the nuclear concentration of hnRNP G-T protein over such a long developmental time window...
might affect splice choices which are critically important for germ cell development. Consistent with dosage of RBMX proteins being important, the RBMX gene family contains two other active paralogues: Rbmx itself and Rbmxrt8. Our expression analysis showed that outside the testis, both Rbmx and Rbmxrt8 (we assume from our above analysis that Rbmxrt14 is not expressed) are both significantly upregulated in the thymus and spleen. Thymus and spleen are important sites for normal immune cell development which requires particularly high levels of alternative splicing (42) which might in turn require increased expression of RBMX proteins. The testis-expressed RBMY gene is also amplified on the mouse and human Y chromosomes, and so might provide increased levels of expression of this protein in spermatogonial cells. Hence, bi-allelic expression levels of Hnrnpgt might be required in the male mouse for normal gene expression pathways during and after meiosis when Rbmx and Rbmxrt8 are not expressed. We postulate that by analogy both copies of HNRNPGT will be also required for normal human germ cell development, and that reductions in hnRNP G-T expression could similarly lead to reduced fertility in men.

**MATERIALS AND METHODS**

**Strain-specific PCR**

The strain-specific PCR primers StrSpecF and StrSpecR (Table 2) span a polymorphic locus and so yield different sized products depending on the strain of mouse. This PCR can distinguish between chromosomal DNA from the mouse strains 129 (ES cell line-derived) and C57Bl6 (blastocyst donor) which contributed to the chimaeras and DBA which is the strain used to isolate uterine sperm. The ratio of PCR products was quantitated (Genetools, SYNGENE) and used to determine the relative contribution of C57Bl/6, DBA and 129 within DNA samples. The uterine sperm samples were collected in DBA mice so that we would be able to differentiate between genomic DNA isolated from the sperm and contaminating genomic DNA from the female.

**Targeted allele-specific PCR**

The targeted allele was detected by a gene-specific PCR in which we utilize the primers TransgenF and TransgenR (Table 2). The primer TransgenF is located upstream of GT but outside the gene targeting vector and the primer TransgenR is located in lacZ. Therefore, a PCR product is only observed if the gene targeting vector is inserted in the correct place in the genome.

**Purification of genomic DNA from sperm**

Sperm were isolated from the uterus or epididymus either in PBS or Universal IVF fluid (Medicult) and purified according to manufacturer’s directions using the Invisorb Spin Forensic kit (Invitrek, Berlin).

**Phylogenetic analysis**

Putative hnRNPGT orthologues and RBMX sequences were identified using both BLAST, and through inspection of the Ensembl and UCSC databases, including multispecies Multiz alignments. Sequences were retrieved from Genbank (human HNRNPGT, accession number NM_014469; mouse Hnrrng, accession number AK005907), bull HNRNPBG, accession number BC114040; mouse Rbmx, accession number NM_011252; human RBMX, accession number NM_002139; human RBMXL1, accession number NM_019610); the Sanger Centre (rat Hnrrng, ENSRNOG00000021946; mouse Rbmxrt8, accession number ENSMUSG00000037070); Mouse Rbmxrt8; Macaque RBMX NM_011252; and ENSEMBL: Tenrec hnrnpbt (ENSETEG00000005619) and Elephant
RBMX (ENSLAFG00000005568); and from UCSC (armadillo and elephant HNRNPGT were both assembled from UCSC vertebrate Multiz Alignments with human chromosome 11 between nucleotides 7066928 and 7068888). Virtual translations of these sequences were aligned using Clustal X (43). Because the open reading frames of the Armadillo and Elephant hnRNPGT sequences were not intact (possibly due to the high throughput nature of the sequence data), these were excluded from the protein alignment, and were manually added to the multiple alignment at the nucleotide level. Trees constructed from the original protein alignment (excluding Armadillo and Elephant hnRNPGT sequences) gave the same two clade topology to that seen in Figure 1 (data not shown). Neighbour joining trees were constructed using the maximum composite likelihood method (44) with a Tamura-Nei model of evolution (45) using a shape parameter of 0.713 estimated from the data using FindModel (http://hcv.lanl.gov/content/hcv-db/findmodel/findmodel.html). All trees were constructed using MEGA4 software (46). Maximum Parsimony and Minimum Evolution trees were also constructed and had very similar topologies (not shown).

**Divergence estimates**

Mean Ks values and standard errors were estimated using the MEGA2.1 software package (47). We used the modified Nei-Gojobori method with Jukes-Cantor correction for multiple substitutions. We set the transition to transversion ratio (R) to 2.5 based on interspecific comparisons of mammalian genes (48). Preliminary alignments of predicted amino acid sequences were used to identify, and exclude from the analysis, regions of uncertain alignment and gaps resulting from deletion or insertion differences. Edited nucleotide sequences were then aligned using Clustal X and imported into MEGA for the analysis. We tested equality of mean Ks values by performing a Z test with type I error equal to 0.05.

**Northern blots**

RNA was purified from mouse tissues using Trizol reagent (Invitrogen, Carlsbad, CA, USA) according to manufacturer’s directions. Fifteen micrograms of total RNA from each tissue was separated by electrophoresis on a 1% formaldehyde gel and blotted onto a Hybond N nylon membrane (Amersham). Because of the high level of sequence identity of Rbmx and Rbmxr8/14 within their open reading frames, we designed probes complementary to their 3′-UTRs. Radiolabelled probes were made from PCR products using either testis cDNA (for Rbmx splice variant 2) or cloned cDNAs (Rbmxspvl1, Rbmxr8/14, Rbm and Hnrnpgt) as a template using the primers shown in Table 2, or from a restriction fragment (Gapdh). Probes were labelled by random priming with (α−32P) dNTPs and were hybridized to membranes as previously described (2). After washing, the membranes were exposed to X-ray film overnight at −80°C next to an intensifying screen. For quantitation, membranes were exposed on the Typhoon (GE healthcare) phosphorimager and band intensities measured and calculated using the Image Quant software. The amount of each transcript relative to the amount of the same transcript in the other tissues was measured to calculate a percentage of the total (if the same amount of transcript was expressed in each tissue then there would be 10% of the total in each of 10 tissues). To correct for loading, we normalized the percentage that we calculated relative to Gapdh (also a percentage) assuming that Gapdh is expressed equally in every tissue (shown numerically on the graph in Fig. 3).

**Raising antisera specific to mouse hnRNPGT and hnRNPG**

We generated and affinity purified a sheep antisera specific to the peptide GGRYDEYRGCSPDGYGG present within the conceptual ORF of mouse hnRNPGT and characterized this as reported previously (2). The detected protein corresponds to mouse hnRN P-G-T since immuno-reactivity was pre-absorbed with the hnRN P-G-T immunising peptide (Supplementary Fig. 2), and it exactly co-migrated with the testis-specific protein detected by the anti-human hnRNPG-T antiserum (2). A peptide antisera against hnRNPG was generated by immunizing rabbits with the peptide RDDGYSTKDC coupled to KLH. Antisera were affinity purified prior to usage.

**Detection of proteins**

Proteins from mouse tissues were homogenized in 2X sample buffer containing 4M urea (49), fractionated by SDS/PAGE and blotted onto a Millipore filter. Western filters were probed with sheep α-hnRNP G-T antiserum (1:500), rabbit α-hnRNPG antiserum (1:500) or rabbit α-Sam68 (Santa Cruz, dilution 1:1000); Rabbit α-ERK2 (Santa Cruz, dilution 1:1000); rabbit α-Elf4E (cell signalling, dilution 1:1000). Antisera were diluted in PBS-tween (0.1%) supplemented with 5% BSA.

**Cell fractionation**

Germ cell purification using elutriation, immunoprecipitations from extracts of purified germ cells and polysome fractionation were carried out as previously described (17).

**Construction of gene targeting vector and generation of chimeraic mice**

Accuprime High Fidelity Taq (Invitrogen) was used to generate the short arm and long arm of the targeting construct using TBV-2 ES cell genomic DNA as a template. The 5′-end containing the short arm was first cloned into plasmid pSKlacZpa so that the coding sequence of Hnrnpgt would be replaced by lacZ. Subsequently, a Kpn1-BamHI fragment containing lacZ and 1.8 kb of sequence upstream of Hnrnpgt was cloned into a modified version of pKSloXPNTSKL (pKSloXPNTSKL + polylinker). Finally, a XhoI-NolI fragment from pNeoflox-8 containing the neo cassette and a 5 kb PacI-NhelI fragment containing the long arm were inserted into the pKSloXPNTSKL + polylinker to yield the final gene targeting vector. The targeting vector was linearized with Kpn1 and electroporated into the ES cells. ES cells were cultured on a layer of SNLP (clone 76/7-4) feeder cells (50). Culture conditions, electroporation, G418 and gancyclovir selection and isolation, propagation and analysis of resistant
ES cell colonies were all as described (51). Resistant ES cell clones were screened by PCR with primers TransgenF and TransgenR (Table 2), and by Southern blot analysis (as shown in Fig. 5). Positive clones were microinjected into C57Bl/6 blastocysts. The resulting chimaeric mice were mated with C57Bl/6 females to screen for germ line transmission, and with DBA females for making uterine sperm samples. Karyotyping of ES cells was carried out by CHROMBOS (Raubling, Germany).

Immunohistochemistry

Tests from the mice and humans were fixed in Bouins solution and embedded in paraffin wax. The sections were prepared according to standard procedures and the antigens were revealed by microwaving in 0.01M citrate buffer (pH6.0) (2). Either a 1:100 dilution of rabbit polyclonal antibody to β galactosidase (Abcam) or an equivalent concentration of rabbit IgG was incubated with the sections overnight for the mouse sections followed by a 1:400 dilution of biotinylated goat anti-rabbit (Dako) secondary antibody. The avidin-biotin-horseradish peroxidise system (ABC-HRP) and dianamobenzidine (Sigma) were utilized for antibody detection. Sections were counterstained in Harris haematoxylin and mounted in Histomount (National Diagnostics). The human testis sections were stained as previously described (11,52) using the antisera specific for hnRNPG and the DAB detection system. For immunofluorescence, the slides were not microwaved before incubation with antibody.

Sperm staining and morphology

Sperm were fixed and stained according to manufacturer’s instructions using Spermac Stain (Stain Enterprises, RSA) and mounted with DPX. Two hundred sperm were counted in each sample and scored either as normal or abnormal based on morphology.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG Online.

ACKNOWLEDGEMENTS

We thank Professor Howard Cooke for helpful advice on the LacZ staining. We would also like to thank Kerstin Seidel for her help with the mice and useful discussions.

Conflict of Interest statement. None declared.

FUNDING

This work was funded by the Biotechnology and Biological Sciences Research Council (BBSRC Grant number G17920 to D.J.E.). C.D. was funded by the Special Trustees of the Royal Victoria Infirmary. The work in C.S.’s laboratory was supported by Teletlon (grant GGP04118 to C.S.) and Lance Armstrong Foundation (to C.S.).

REFERENCES


