 Human Molecular Genetics, 2000, Vol. 9, No. 2 311–319

A human sex-chromosomal gene family expressed in male germ cells and encoding variably charged proteins

Bruce T. Lahn+ and David C. Page

Howard Hughes Medical Institute, Whitehead Institute and Department of Biology, Massachusetts Institute of Technology, 9 Cambridge Center, Cambridge, MA 02142, USA

Received 1 October 1999; Revised and Accepted 5 November 1999 DDBJ/EMBL/GenBank accession nos AF159127–AF159129 and AF000979

Approximately 12 X-Y homologous gene pairs have been identified in the non-recombining portions of human sex chromosomes. These X-Y gene pairs fall into two categories. In the first category, both X and Y homologs are ubiquitously expressed. In the second category, the X homolog is ubiquitously expressed, whereas the Y homolog is expressed exclusively in the testis. Here we describe a family of human X-Y genes that cannot be assigned to either category. Designated *VCX/Y* (*Variable Charge X/Y*; *VCY* previously known as *BPY1*), this gene family has multiple members on both X and Y, and all appear to be expressed exclusively in male germ cells. Members of the *VCX/Y* family share a high degree of sequence identity, with the exception that a 30 nucleotide unit is tandemly repeated in X-linked members but is present only once in Y-linked members. These atypical features suggest that the *VCX/Y* family has evolved in a manner previously unrecognized for mammalian X-Y genes. We also found that a copy of *VCX* is present in CRI-S232, a previously described genomic fragment derived from the X chromosome. Studies have shown that aberrant recombination between arrays of CRI-S232-homologous repeats flanking the *steroid sulfatase* (*STS*) gene results in *STS* deletion, which is manifested clinically as X-linked ichthyosis. The revelation that CRI-S232 contains *VCX* offers a more precise description of the genetic etiology of X-linked ichthyosis: it results from aberrant recombination between *VCX* gene arrays that flank the *STS* locus.

INTRODUCTION

Mammalian sex chromosomes evolved from a pair of autosomes (1–4). An irreversible process during sex chromosome evolution is the suppression of X-Y recombination over progressively larger regions (5). This process affected the two sex chromosomes in radically different ways. Within the non-recombining portion of the X (NRX; the portion that does not recombine with the Y), genes remain well preserved. In contrast, within the non-recombining portion of the Y (NRY), most genes have degenerated (5,6). There are, however, exceptions to the general trend of Y chromosome degeneration: a handful of ancestral genes were found to have persisted in the NRY as well as the NRX of extant mammals, where they exist as differentiated homologs (7). Prior studies of these X-Y homologous genes have shown that two adaptive processes may have been responsible for the persistence of their Y-linked members. The first is the conservation of certain essential housekeeping genes on both X and Y chromosomes to ensure double dosage of these genes in males and females (7). Y-linked genes preserved by this process resemble their X homologs in that they encode widely distributed housekeeping proteins. In addition, their X homologs escape X-inactivation to fulfill the double dosage requirement (7,8). The second process is the selection for, and subsequent preservation of, genes on the Y that have acquired male-beneficial functions (9,10). Y-linked genes preserved by this process are distinct from their widely expressed X homologs in that they are expressed only in the testis (9–11).

We had previously identified a single cDNA sequence corresponding to the *BPY1* (*Basic Protein Y 1*) gene(s) on the human Y chromosome (7). For reasons that will become obvious in later text, this gene is now renamed *VCY* for *Variable Charge Y*. *VCY* is expressed only in the testis, and encodes a small, positively charged protein of unknown function. We initially thought that, like most other testis-specific genes on the human Y chromosome, *VCY* lacked X homologs (7). This conclusion was overturned when we isolated additional cDNA clones using *VCY* cDNA as a probe, and showed that many of these clones derive from close homologs of *VCY* on the human X chromosome (named *VCX* for *Variable Charge X*). Expression analysis showed, to our surprise, that all copies of *VCX* and *VCY* are transcribed exclusively in the testis, most likely in male germ cells. This feature distinguishes this X-Y gene family from the two previously recognized categories of X-Y genes. Models are proposed for the evolution of this gene family on human sex chromosomes.

+To whom correspondence should be addressed at present address: Department of Human Genetics, University of Chicago, 924 East 57th Street, Chicago, IL 60637, USA. Tel: +1 773 834 4393; Fax: +1 773 834 0505: Email: blahn@genetics-uchicago.edu

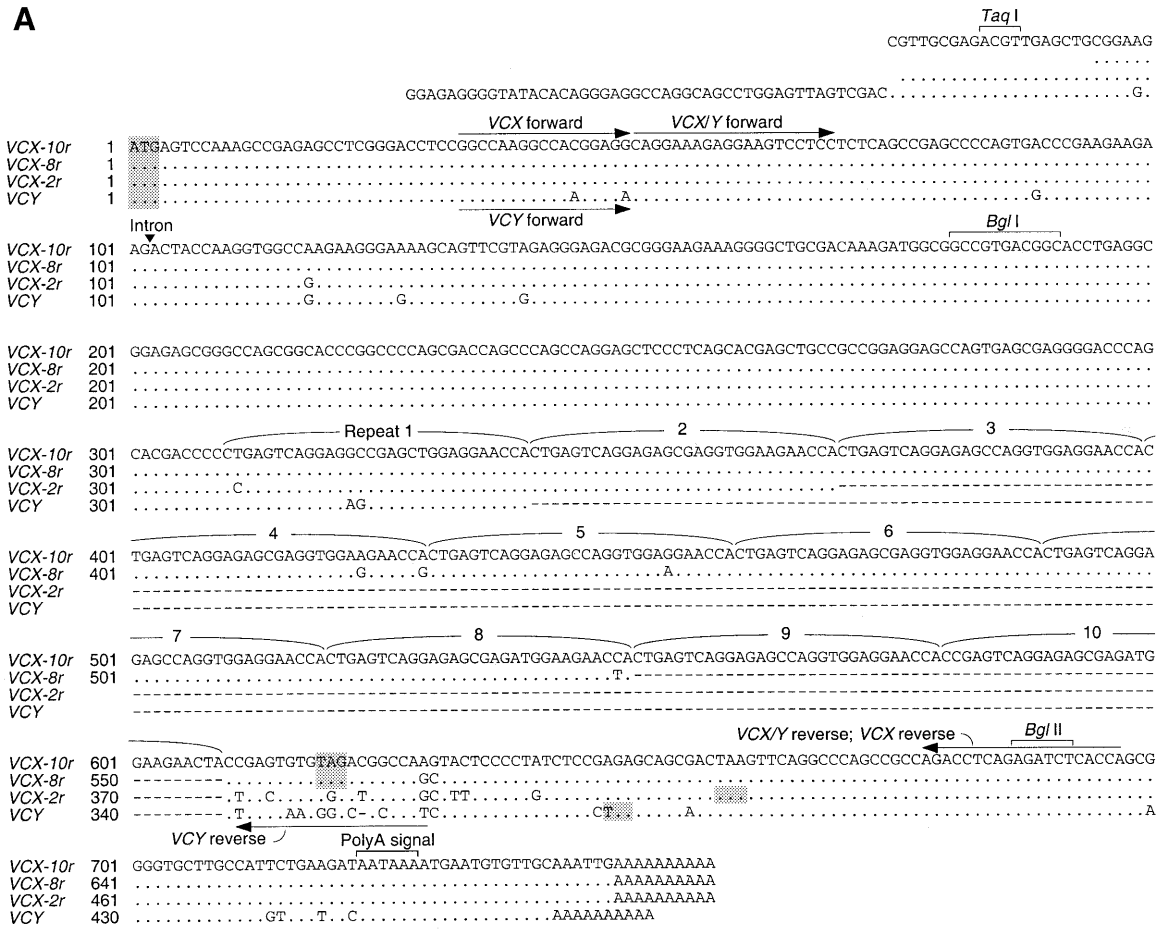

Figure 1. Alignment of nucleotide and deduced protein sequences of the four cDNA species. Dot, identity to consensus; dash, absence of base or amino acid. (**A**) Nucleotide alignment of Y-linked *VCY* (previously known as *BPY1*) and its X-linked homologs that contain 2 (*VCX-2r*), 8 (*VCX-8r*) or 10 (*VCX-10r*) of the 30 nucleotide repeat units. Repeats are numbered. Start and stop codons are shaded. The single intron, the polyadenylation signal, and several restriction sites (*Taq*I, *Bgl*I and *Bgl*II) are indicated. Locations of primers used for radiation hybrid mapping of *VCX*, and for RT–PCR expression analysis of *VCX* and *VCY*, are indicated by arrows (see Fig. 5 for details of the RT−PCR analysis).

RESULTS

When the previously isolated *VCY* cDNA clone was used to probe human genomic Southern blots, complex banding patterns were observed in both males and females, suggesting the existence of *VCY*-homologous sequences outside the Y chromosome. When the same probe was used to screen a human testis cDNA library, 14 clones were isolated. These 14 clones make up four distinct cDNA species (Fig. 1). One cDNA species (represented by five clones) is identical in sequence to *VCY* as previously reported. The remaining three cDNA species are distinguishable by the fact that a 30 nucleotide unit is tandemly repeated twice (six clones), eight times (one clone) or 10 times (three clones) within each cDNA sequence. This repeat unit is present only once in *VCY*. Outside the repeat regions, the four cDNA species share 96% or greater pairwise nucleotide identity within their open reading frames. PCR amplifications of genomic DNAs using cDNA-based primers revealed that there is a single intron, 192 nucleotides in length, in *VCY* and its homologs (indicated in Fig. 1A).

Additional studies (detailed below) showed that the cDNA species with one internal repeat unit corresponds to two identical copies of *VCY* on the Y chromosome, and that the cDNA species with two or more internal repeats correspond to multiple homologs of *VCY* on the X chromosome (named *VCX*). To distinguish between *VCX* genes with different numbers of repeats, an appendix will be used to indicate repeat number. The *VCX* gene(s) with two repeats, for example, will be referred to as *VCX-2r*.

We speculated that additional *VCX* or *VCY* genes might exist, carrying repeat arrays different in length from those represented by the cDNA clones that we had isolated. To address this possibility, Southern blots were generated from *Bgl*I–*Bgl*II double-digested genomic DNAs of unrelated individuals, and probed with a fragment from the previously isolated *VCY* cDNA. *Bgl*I and *Bgl*II were chosen because their restriction sites (indicated in Fig. 1A) flank the repeat regions in all four cDNA species. Numerous bands were detected in each individual (Fig. 2). One band is apparently Y-linked, present only in males. This band corresponds in length to *VCY*, which contains only one internal repeat unit. The other bands

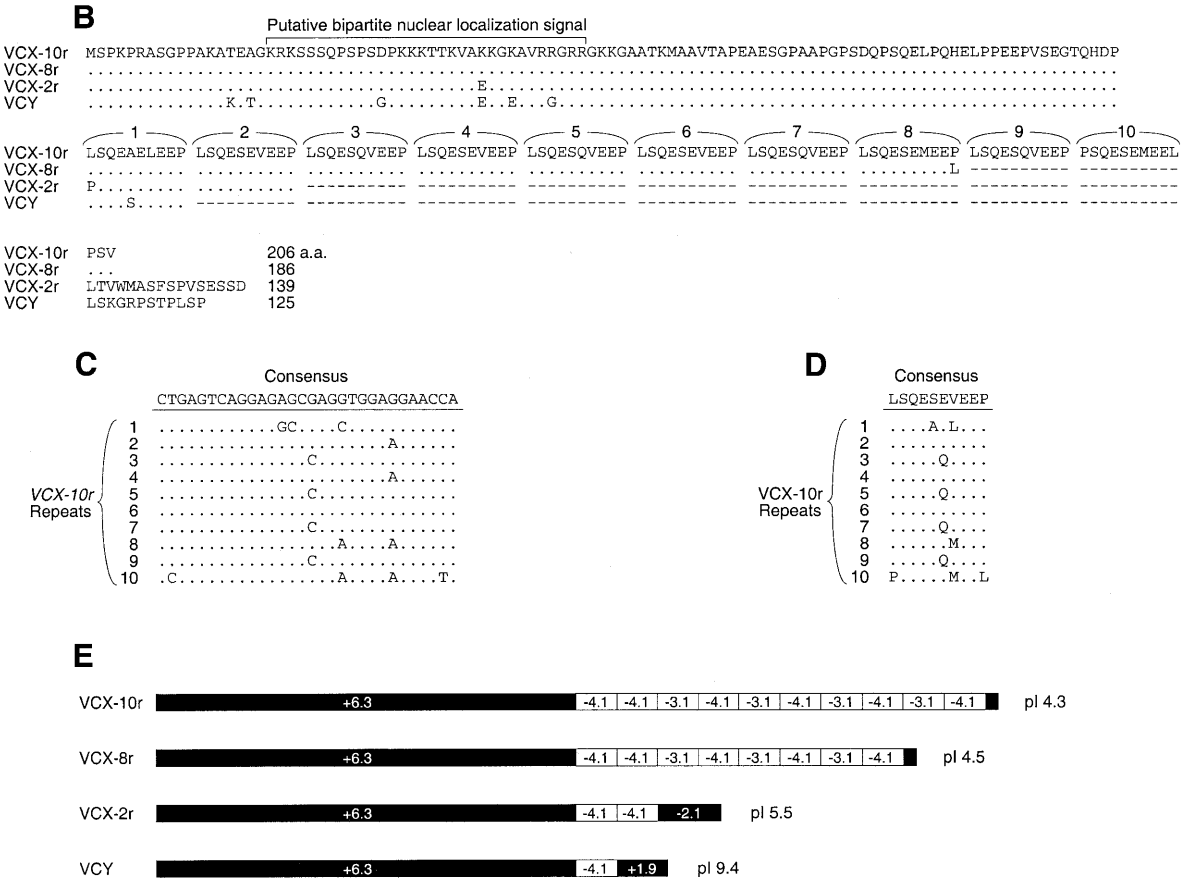

Figure 1. Alignment of nucleotide and deduced protein sequences of the four cDNA species. Dot, identity to consensus; dash, absence of base or amino acid. (**B**) Protein alignment of VCY, VCX-2r, VCX-8r and VCX-10r. Repeats are numbered. The putative bipartite nuclear localization signal is indicated. (**C**) Nucleotide alignment of the 10 repeats of *VCX-10r*. Repeats 1 and 10 are the most divergent; repeats 2–9 alternate imperfectly between two sequence types. (**D**) Protein alignment of the 10 repeats of VCX-10r. Like nucleotide sequence, repeats 2–9 alternate between two sequence types. (**E**) Charge diagrams of the four protein species. Open bar, the 10 a.a. repeat motif; solid bar, non-repeat region. The amount of calculated charge (at pH 7) carried by each segment of the proteins is indicated. The calculated isoelectric point (pI) of each protein is indicated.

correspond in length to genes that contain two or more repeats. The two- and eight-repeat-containing bands are present in all males and females tested. The remaining bands, which contain >8 and up to 30 internal repeats, are highly polymorphic, and do not appear to be male-specific.

To obtain meiotic segregation patterns of the *VCX* and *VCY* genes, Southern blots were generated from *Taq*I-digested genomic DNA samples of two unrelated three-generation kindreds, and probed with the same *VCY* cDNA fragment. Again, numerous bands were detected in each individual (Fig. 3). Two monomorphic bands show Y linkage, as they are present only in males, and are passed consistently from father to son. The remaining bands are polymorphic and are apparently X linked, passed consistently from father to daughter, but never from father to son. These X-specific bands are tightly linked to each other, segregating as a single locus in all cases except one, where a recombination event occurred between the two maternal haplotypes (indicated in Fig. 3).

One apparent contradiction between the *Bgl*I–*Bgl*II and the *Taq*I Southern analyses is the presence of a single male-specific band corresponding to *VCY* in the former, but two male-specific bands in the latter. We speculated that two copies of *VCY* might exist on the Y chromosome (distinguishable by *Taq*I digest), each of which contains one repeat unit (therefore not distinguishable by *Bgl*I–*Bgl*II double digest). To investigate this possibility, six Y chromosomal BAC clones were isolated by probing a genomic BAC library with the *VCY* cDNA fragment. Southern blots of the six clones prepared either by *Bgl*I–*Bgl*II double digest, or by *Taq*I single digest were probed with the same *VCY* cDNA fragment (Fig. 4). In the case of the *Bgl*I–*Bgl*II double digest, a single band was observed for all six clones. This band is identical in size to the male-specific *VCY* band in the *Bgl*I–*Bgl*II genomic Southern analysis (compare Fig. 4 with Fig. 2), indicating that all six clones contain *VCY*. In the case of the *Taq*I digest, two bands of different sizes were observed, corresponding in size to the two male-specific bands in the *Taq*I genomic Southern analysis (compare Fig. 4 with Fig. 3). Five BAC clones contain either of the two bands; one clone contains both bands. These results confirm the presence of two copies of *VCY* on the Y chromosome, which are separated by <140 kb—the length of the BAC clone that contains both copies of *VCY* (the size of

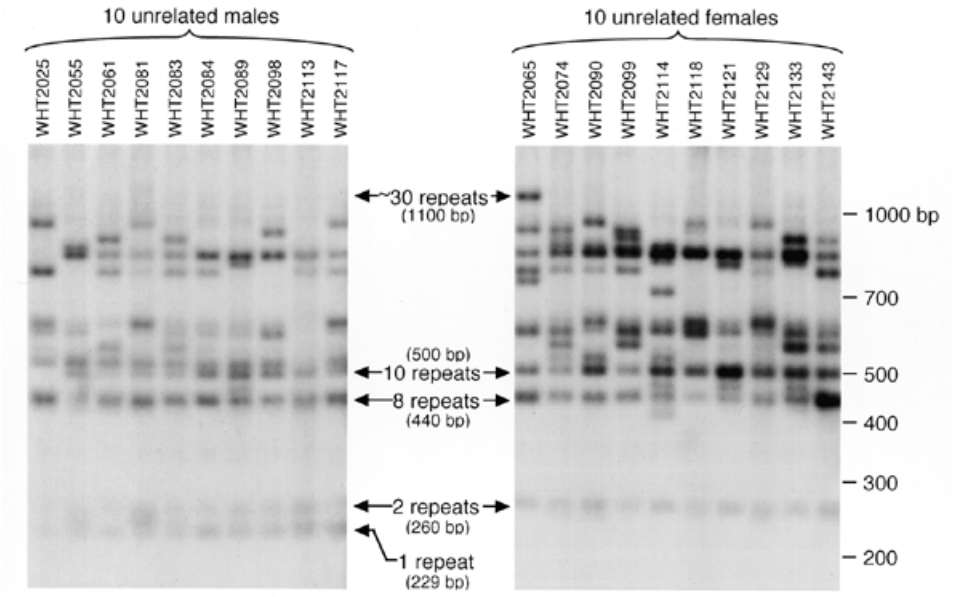

Figure 2. Southern blot detection of *VCY* homologous sequences. Genomic DNA samples from unrelated males and females were double digested with *Bgl*I and *Bgl*II, and probed with a *VCY* cDNA fragment that lies between *Bgl*I and *Bgl*II sites (Fig. 1A). The estimated number of repeats (based on 30 bp/repeat) within each *VCY* homologous fragment is indicated.

this BAC clone was estimated by restriction fingerprinting). Partial sequencing of the BAC clones revealed complete nucleotide identity between open reading frames of the two *VCY* copies. The fact that a single *VCY* band in the *Bgl*I–*Bgl*II Southern blot corresponds to two copies of the gene raises the possibility that other bands in the Southern blot may also represent multiple gene copies that each have the same number of internal repeats. By counting the number of X-specific bands in the *Bgl*I–*Bgl*II genomic Southern analysis, and taking into consideration that more intense bands may represent multiple copies (Fig. 2), we estimated that ~12 *VCX* genes may be present on a typical X chromosome.

We had previously mapped *VCY* to Y chromosome deletion interval 5G (7). By designing primers specific to *VCX* sequences (locations of primers indicated in Fig. 1A) and typing them on a radiation hybrid (RH) panel (12), we localized *VCX* to the same RH map position as the *steroid sulfatase* (*STS*) locus on distal Xp, corresponding to cytogenetic band Xp22.3 (see ref. 5 for a map of the X chromosome that depicts the *STS* locus).

We conclude that the *VCX* and *VCY* genes constitute a large family. Two copies are located no more than 140 kb from each other on the Y. The remaining dozen or so copies are located in close proximity on the X. A 30 nucleotide unit is present once in the two Y-linked copies, but is tandemly repeated two or more times in all the X-linked copies. Amplification of the X-linked copies has occurred at both the level of entire genes and the level of the 30 nucleotide repeat unit within each gene.

The gene family encodes proteins with variable charge

The 30 nucleotide repeat unit in the *VCX*/*Y* gene family encodes a 10 amino acid motif that is rich in the acidic residue Glu, and is predicted to be highly negatively charged. Outside the repeat regions, the VCX/Y proteins are predicted to be highly positively charged, owing to an abundance of the basic residues Arg and Lys (Fig. 1B). Therefore, the number of repeats in a given protein should, in theory, exert a strong influence on its overall charge. The Y-encoded VCY proteins contain only one repeat motif. Their overall charge is predicted to be highly positive, with a calculated isoelectric point (pI) of 9.4. In contrast, the X-encoded VCX proteins contain two or more repeats. Their overall calculated charge ranges from moderately negative (pI 5.5 in the case of two repeats), to highly negative (pI 4.3 in the case of 10 repeats) (Fig. 1E). Sequence variations at 3′ ends of *VCX*/*Y* coding regions (where there is a high concentration of nucleotide substitutions amongst the various *VCX*/*Y* cDNA sequences) also contribute to the charge differential between X-encoded and Y-encoded proteins (Fig. 1A and E). This charge variability is the rationale for referring to these genes as *Variable Charge X* and *Variable Charge Y*.

Two features of the deduced VCX/Y proteins—their small size and high charge—resemble those of chromatin-associated proteins such as histones and HMG proteins. Based on this resemblance, we had previously speculated that VCY might interact with nucleic acids (7). Motif searches against the PROSITE and SWISS-PROT databases for protein families and domains identified a putative bipartite nuclear localization signal near the N-terminus of VCX/Y (Fig. 1B), suggesting that they are nuclear proteins. (A description of bipartite nuclear localization signals is available at http://www.expasy.ch/cgi-bin/get-prodoc-entry?PDOC00015 .) Searches for structural patterns failed to identify any folded motifs (i.e. α-helices or β-sheets) in these proteins, suggesting that they are unlikely to fold into stable structures on their own without interacting with other cellular components such as nucleic acids or protein factors (13). These features are consistent with, but provide no direct evidence for, the VCX/Y proteins being components of chromatin.

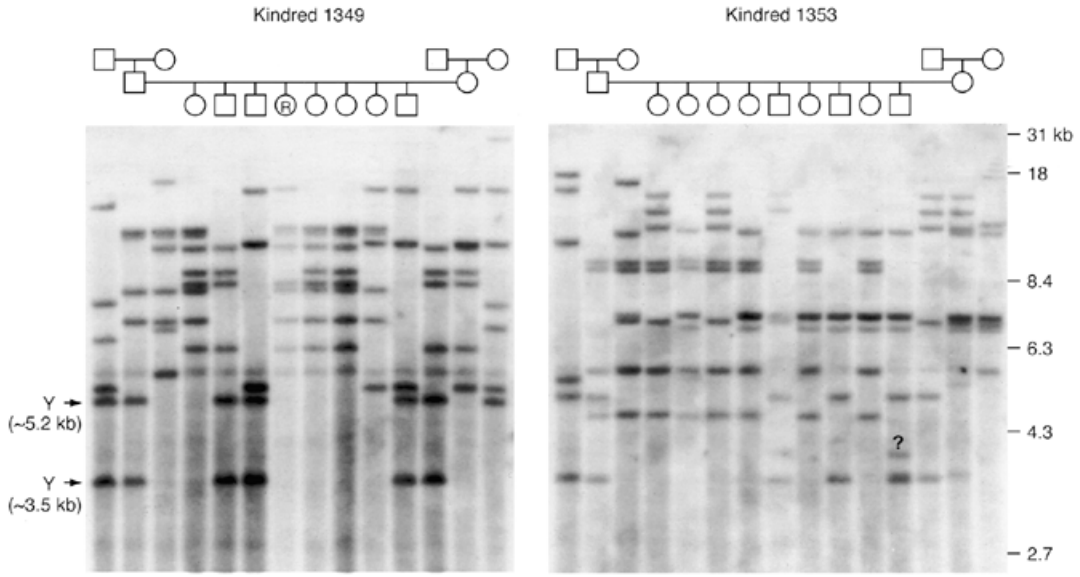

Figure 3. Southern blot analysis of segregation patterns of *VCY* homologous fragments in two kindreds. Genomic DNAs were digested with *Taq*I, and probed with the same *VCY* cDNA fragment as used in Figure 2. Y-specific bands (present only in males) are indicated. In most cases, the parental haplotype is transmitted without recombination. In one female (R), recombination between the two maternal haplotypes has occurred. One individual carries a fragment that is not present in either of his parents (indicated by a question mark). This fragment may be the result of a new mutation.

Expression of the *VCX/Y* family is likely restricted to male germ cells

We had previously probed a multiple-tissue northern blot with a *VCY* cDNA fragment, and observed signal only in the testis (7). Since the probe did not distinguish among the various members of the *VCX/Y* family, the northern data suggested that expression of the entire family is restricted to the testis. We confirmed the testis-specific expression of *VCX/Y* by performing multiple-tissue RT–PCR using primers specific for *VCX*, *VCY* or both (Fig. 5A).

To examine whether the family is expressed in germ cells or in the somatic portion of the testis, we performed RT–PCR on a biopsy sample of a testis that lacked germ cells (Sertoli cells only). In this case, *VCX/Y* transcripts were not detected, suggesting that expression of the family is restricted to male germ cells (Fig. 5B). However, our data cannot rule out the possibility that the *VCX/Y* family is in fact expressed in the somatic portion of the testis, and that the failure to detect *VCX/Y* transcript in the germ cell-deficient testis is due to the somatic expression being dependent on the presence of germ cells.

For most X-chromosomal genes that have Y-linked homologs, it is of interest to address their X-inactivation status in female cells. But in the case of the *VCX* genes, we were unable to address this issue since the only place where these genes are detectably expressed is the testis, a tissue present only in males.

A copy of *VCX* is present in the genomic clone CRI-S232

The repetitive nature of the *VCX/Y* family on sex chromosomes and the close proximity of the *VCX* genes to the X-linked *STS* locus are reminiscent of the CRI-S232 homologous sequences on human sex chromosomes (14). CRI-S232 is a previously isolated anonymous human genomic clone. When CRI-S232 was used to probe genomic Southern blots, it detected multiple polymorphic fragments that mapped to distal Xp, adjacent to the *STS* locus, as well as a set of monomorphic fragments that mapped to proximal Yq (14). The CRI-S232 clone has been restriction mapped and partially sequenced (15). Comparison of the CRI-S232 partial sequence with *VCX/Y* cDNA sequences revealed that CRI-S232 contains a copy of *VCX* (Fig. 6).

It has been shown that CRI-S232-homologous sequences on the X chromosome reside on both sides of *STS*, and that recombination between these flanking sequences accounts for the majority of *STS* deletions (16). The clinical manifestation of *STS* deletion is ichthyosis, or scaly skin syndrome (17). The common occurrence of this aberrant recombination has resulted in a high allele frequency of *STS* deletion (~0.01%) in the population (16). Our finding that CRI-S232 contains a copy of *VCX* offers a more precise description of the molecular etiology of *STS* deletions, namely that this frequent genetic defect is the result of aberrant recombination between *VCX* gene arrays flanking the *STS* locus.

DISCUSSION

The process of Y degeneration is so effective during sex chromosome evolution that only those genes with particular functional characteristics could persist. Previous studies of X-Y homologous genes within human NRX and NRY have shown that the persistence of their Y-linked members is the result of two adaptive processes: (i) conservation of certain ubiquitously expressed housekeeping genes (7); or (ii) evolution of male-specific function and testis-specific expression (9,10). Most X-homologous genes on the human Y chromosome have apparently been maintained through the former process, as both X and Y homologs of these genes encode ubiquitously expressed housekeeping proteins (7). Only two X-homologous genes on the human Y are known to have been preserved

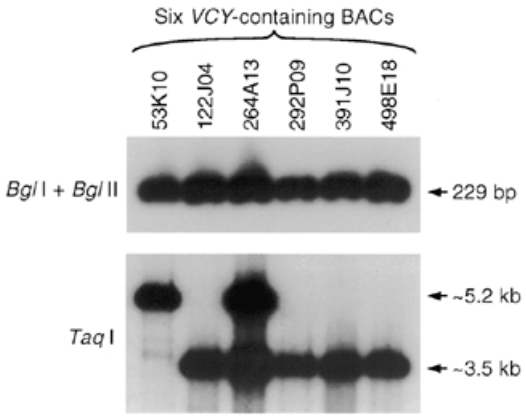

Figure 4. Southern blot analysis of six *VCY*-containing genomic BAC clones. The clones were either double digested with *Bgl*I and *Bgl*II, or digested with *Taq*I, and probed with the same *VCY* cDNA fragment as in Figures 2 and 3.

through the process of male specialization: *SRY*, the male-determining gene (18), and *RBMY*, a putative spermatogenic factor (19). Both *SRY* and *RBMY* have evolved male-specific functions, whereas their homologs on the X, *SOX3* and *RBMX*, respectively, retain broad patterns of expression that presumably resemble the ancestral expression status (9–11). *RBMY* has undergone amplification (19), which is a hallmark of Y genes expressed exclusively in the testis. Some of these testis-specific genes arrived on the Y by transposition or retroposition of autosomal genes, rather than by persistence (20,21). In the mouse, there are three additional examples—*Zfx*/*y*, *Ube1x*/*y* and *Dffrx*/*y*—of X-Y genes that have become testis-specific on the Y, while maintaining ubiquitous expression on the X (22–26). Like *RBMY* in humans, *Zfy* and *Ube1y* in the mouse have undergone amplification on the Y chromosome (22,24,27).

Certain features of *VCY* resemble those of *RBMY*: both *VCY* and *RBMY* are expressed exclusively in the testis and have undergone amplification on the Y chromosome. But the resemblance does not extend to their X homologs. *VCX*, like *VCY*, is expressed only in the testis, and has undergone amplification. *RBMX* on the other hand, is widely expressed in somatic tissues and is present in a single copy (28).

What then is the adaptive process that contributed to the preservation of *VCY* on the human Y chromosome? In the absence of additional data on the function of the *VCX*/*Y* gene family, we will offer two speculative models: (i) the teamwork model; and (ii) the selfish-gene model.

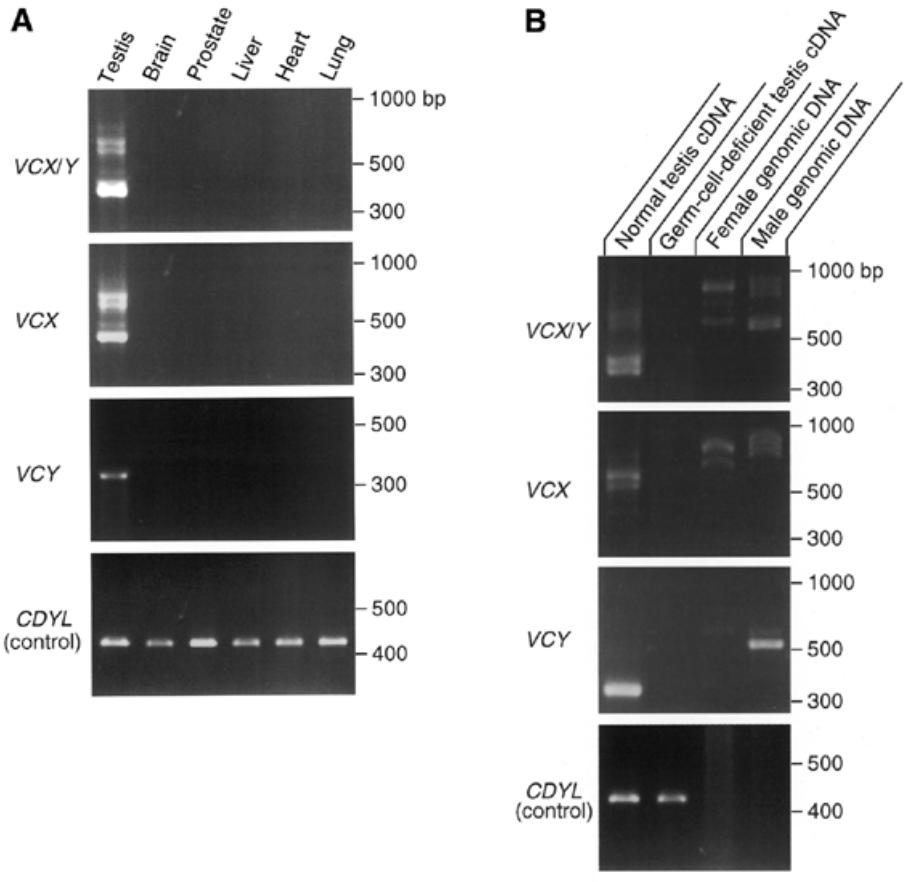

Figure 5. RT–PCR analysis of *VCX*/*Y* expression. Four separate PCR assays were performed on the cDNA of each tissue. Locations of primers in the *VCX* and *VCY* genes are indicated in Figure 1A. *VCX/Y*, primers that amplify both *VCX* and *VCY*; *VCX*, primers that amplify *VCX* only; *VCY*, primers that amplify *VCY* only; *CDYL*, positive control primers that amplify the ubiquitously transcribed *CDYL* (21). (**A**) Expression in multiple tissues of normal males. (**B**) Expression in normal versus germ cell-deficient testis. The apparent banding pattern difference between (A) and (B) in the normal testis lane is due to polymorphism of the internal repeat array sizes in the *VCX* genes between the two cDNA samples (which are derived from two individuals).

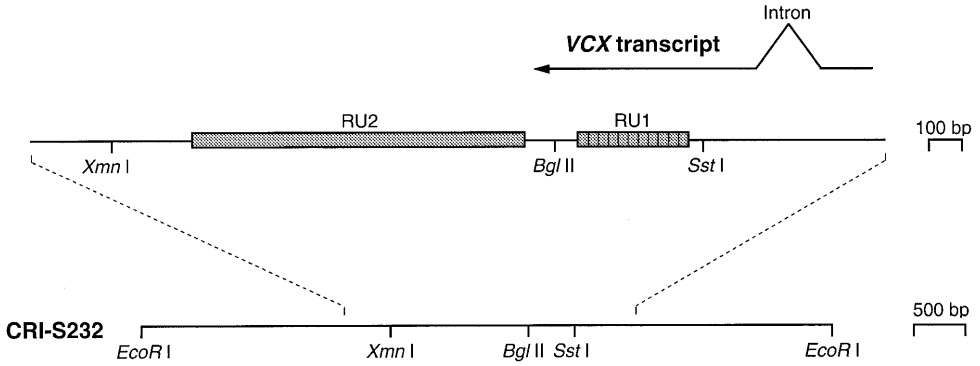

Figure 6. Organization of CRI-S232 and its resident copy of *VCX*. The portion of the diagram that depicts CRI-S232 genomic structure is adapted from a figure by Knowlton *et al.* (14). Two sequence elements, RU1 and RU2 as referred to by Knowlton *et al.* (14), are shaded. RU1 corresponds to the tandem array found in the *VCX* transcripts. It contains 11 copies of the 30 nucleotide repeat unit. RU2 is a low-complexity sequence element (14). A putative *VCX* transcript is drawn above its corresponding genomic region.

According to the teamwork model, various members of the VCX/Y protein family complement each other in function to collectively mediate a certain process in spermatogenesis. By this model, the *VCY* genes have been preserved during evolution because their spermatogenic functions are somewhat distinct from those of the *VCX* genes.

The selfish-gene model is inspired by previous studies of sex chromosome meiotic drive in certain insect species. In these species, 'selfish' genes are believed to have evolved first on the X chromosome, causing the X to be transmitted more often than the Y. The Y chromosome counters by evolving suppressors of the X-linked selfish genes (29–33). In *Drosophila melanogaster*, two sex chromosome loci—*Stellate* on the X and *crystal* on the Y—have been implicated in sex chromosome meiotic drive, and are thought to be selfish genetic elements (31,32). The *VCX/Y* gene family resembles *Stellate*/*crystal* in that it is amplified on both sex chromosomes, with X- and Y-linked members active only in male germ cells. Such resemblance, together with the extreme charge differential between X- and Y-encoded proteins, raises the tantalizing possibility that *VCX* and *VCY* are selfish genes, wherein X-encoded proteins (which are highly negatively charged) and Y-encoded proteins (which are highly positively charged) antagonize each other in a race to distort the transmission ratio of the sex chromosomes to their advantage. In humans, slightly more males are conceived than females (34). If the *VCX/Y* gene family members are indeed selfish genetic elements as we speculated, they may play a role in such sex ratio distortion.

The *VCX/Y* gene family has been evolving rapidly on human sex chromosomes, as shown by its poor conservation in mammals. Low-stringency hybridization to zoo-blots with the CRI-S232 clone (which contains a copy of *VCX*) has previously shown that CRI-S232 homologous sequences are detectable only in simian primates, but not in prosimians or non-primate mammals (16). This result is consistent with our inability to isolate mouse homologs of *VCX/Y* despite repeated attempts. The lack of conservation of the *VCX/Y* family reflects either rapid sequence change of ancient genes, or recent *de novo* emergence of this gene family in the simian lineage.

Members of the *VCX/Y* family share a high degree of sequence similarity. Of the four *VCX/Y* cDNA species that we identified, the three *VCX* species share 99% pairwise nucleotide identity within their open reading frames (not considering the internal repeats) (Fig. 1A). Partial sequencing of *VCY* BAC clones showed that the two copies of *VCY* share complete identity in their open reading frames. Between *VCX* and *VCY*, sequence similarity is slightly lower, ~96%. These observations indicate that the multiple members of the *VCX/Y* family derived from a single sequence very recently during evolution through gene amplification, and possibly, subsequent gene conversion amongst amplified copies. They also suggest that the amplification of *VCX* on the X and *VCY* on the Y occurred after the divergence of X- and Y-linked sequences.

Close examination of nucleotide substitutions between *VCX* and *VCY* suggest that sequences of this gene family have undergone positive selection. All 13 substitutions between the *VCX* and *VCY* coding regions (>348 bases; not considering the internal repeats) are non-synonymous (they alter amino acids). Given that ~25% of random nucleotide substitutions should be silent (35), the probability that all 13 substitutions would be non-silent is only 2.4%. More likely, these substitutions are the result of positive selection for novel protein sequences with altered cellular functions.

X-linked members of the *VCX/Y* family exhibit an extremely high degree of polymorphism in the lengths of the internal tandem arrays present in each gene (Fig. 2). This high degree of polymorphism is most likely the result of rapid and continual expansions and contractions of individual arrays within each gene, which may be due to uneven crossovers between homologous chromosomes or sister chromatids. Such uneven crossovers may also produce polymorphism in the copy number of *VCX*, a prediction that we did not test since the precise copy number of *VCX* was difficult to ascertain.

It is rare that a gene family possesses so many atypical properties as does *VCX/Y*. The stage is now set for further investigation of this gene family, especially at the level of protein function, to better understand its precise role in spermatogenesis, and its evolution on human sex chromosomes.

MATERIALS AND METHODS

Identification of cDNA clones

The *Bgl*I and *Bgl*II flanked region (Fig. 1A) of the previously isolated *VCY* (previously known as *BPY1*) cDNA clone (7) was labeled with [^{32}P]dCTP by random priming, and used to screen a human adult testis cDNA λ phage library (Clontech, Palo Alto, CA). Library blots were incubated with probe overnight at 65°C in 0.5 M Na_iPO_4 (pH 7), 7% sodium dodecyl sulfate (SDS), followed by three washes of 15 min each at 65°C in 0.1× standard saline citrate, 0.1% SDS.

Identification of BAC clones

RPCI-11 human BAC library filters (36) (Roswell Park Cancer Institute, Buffalo, NY) were screened with the same probe and conditions as that used for identifying *VCY*-homologous cDNA clones. Positive clones were further screened by a previously described *VCY*-specific PCR assay (7) to identify *VCY* clones. Clone numbers for the six BACs are listed in Figure 4.

Southern blot analyses

For genomic Southern blots, each lane contained 7 µg of genomic DNA. For Southern blotting of the BAC clones, each lane contained 5 ng of BAC DNA. Hybridization probe and conditions were the same as those used for identifying *VCY* homologous cDNA clones.

Radiation hybrid mapping of *VCX*

Using PCR, 93 RH cell lines of the GeneBridge 4 panel (12) (Research Genetics, Huntsville, AL) were assayed for the presence of *VCX*. Locations of PCR primers in *VCX* are indicated in Figure 1A; their sequences are GGCCAAGGCCACGGAGG and TGGTGAGATCTCTGAGGTCT. Thermocycling conditions: 30 cycles of 1 min at 94°C, 45 s at 60°C, and 45 s at 72°C. Analysis of the PCR results positioned *VCX* with respect to the RH map constructed at the Whitehead/MIT Center for Genome Research (37) (http://www-genome.wi.mit.edu/cgi-bin/contig/phys_map). *VCX* had exactly the same RH map location as the *STS* gene.

Protein pattern and profile searches

Searches for protein motifs were performed against the PROSITE and SWISS-PROT databases using web-based search engines ScanProsite and ProfileScan. Both search engines can be found at http://www.expasy.ch/prosite/ .

RT–PCR analysis of *VCX*/*Y* expression

Testis, brain, prostate, liver and heart cDNA samples from normal males were purchased from Clontech. To make cDNA from the testis biopsy sample, total RNA was first extracted using Trizol reagent (Gibco BRL, Gaithersburg, MD), followed by cDNA synthesis using the Advantage kit (Clontech). Locations of PCR primers used to generate the data shown in Figure 5 are indicated in Figure 1A; their sequences are as follows:
VCX/*Y*, CAGGAAAGAGGAAGTCCTCC
and TGGTGAGATCTCTGAGGTCT;
VCX, GGCCAAGGCCACGGAGG
and TGGTGAGATCTCTGAGGTCT;
VCY, GGCCAAGGCCAAGGAGA
and ATGGGCGCCCCTTACTCA;
CDYL, GTACATCTCCGTTCATGGATG
and CTGATAGCTTCTGCCATTTTAG.
All primer pairs spanned introns. PCR conditions were the same as that used for radiation hybrid mapping of *VCX*.

GenBank accession numbers

GenBank accession numbers for *VCX* and *VCY* cDNA sequences are as follows: *VCX-2r*, AF159127; *VCX-8r*, AF159128; *VCX-10r*, AF159129; *VCY*, AF000979. The *VCY* cDNA sequence was previously published (7).

ACKNOWLEDGEMENTS

We thank R. Oates for patient samples; H. Skaletsky for database searches and sequence analysis; L. Brown, T. Kawaguchi and R. Saxena for technical advice; and J. Berger, P. Carr and S. Rozen for discussions and comments on the manuscript. This work was supported by the National Institutes of Health.

REFERENCES

1. Bull, J.J. (1983) *Evolution of Sex Determining Mechanisms.* Benjamin Cummings, Menlo Park, CA.
2. Graves, J.A. (1996) Mammals that break the rules: genetics of marsupials and monotremes. *Annu. Rev. Genet.*, **30**, 233–260.
3. Charlesworth, B. (1996) The evolution of chromosomal sex determination and dosage compensation. *Curr. Biol.*, **6**, 149–162.
4. Rice, W.R. (1996) Evolution of the Y sex chromosome in animals. *BioScience*, **46**, 331–343.
5. Lahn, B.T. and Page, D.C. (1999) Four evolutionary strata on the human X chromosome. *Science*, **286**, 964–967.
6. Jegalian, K. and Page, D.C. (1998) A proposed path by which genes common to mammalian X and Y chromosomes evolve to become X inactivated. *Nature*, **394**, 776–780.
7. Lahn, B.T. and Page, D.C. (1997) Functional coherence of the human Y chromosome. *Science*, **278**, 675–680.
8. Brown, C.J., Carrel, L. and Willard, H.F. (1997) Expression of genes from the human active and inactive X chromosomes. *Am. J. Hum. Genet.*, **60**, 1333–1343.
9. Delbridge, M.L., Lingenfelter, P.A., Disteche, C.M. and Graves, J.A. (1999) The candidate spermatogenesis gene RBMY has a homologue on the human X chromosome. *Nature Genet.*, **22**, 223–224.
10. Mazeyrat, S., Saut, N., Mattei, M.-G. and Mitchell, M.J. (1999) RBMY evolved on the Y chromosome from a ubiquitously transcribed X-Y identical gene. *Nature Genet.*, **22**, 224–226.
11. Foster, J.W. and Graves, J.A. (1994) An SRY-related sequence on the marsupial X chromosome: implications for the evolution of the mammalian testis-determining gene. *Proc. Natl Acad. Sci. USA*, **91**, 1927–1931.
12. Gyapay, G., Schmitt, K., Fizames, C., Jones, H., Vega-Czarny, N., Spillett, D., Muselet, D., Prud'Homme, J.F., Dib, C., Auffray, C. *et al.* (1996) A radiation hybrid map of the human genome. *Hum. Mol. Genet.*, **5**, 339–346.
13. Creighton, T.E. (1993) *Proteins: Structures and Molecular Properties.* W.H. Freeman and Co., New York, NY.
14. Knowlton, R.G., Nelson, C.A., Brown, V.A., Page, D.C. and Donis-Keller, H. (1989) An extremely polymorphic locus on the short arm of the human X chromosome with homology to the long arm of the Y chromosome. *Nucleic Acids Res.*, **17**, 423–437.
15. Li, X.M., Yen, P.H. and Shapiro, L.J. (1992) Characterization of a low copy repetitive element S232 involved in the generation of frequent deletions of the distal short arm of the human X chromosome. *Nucleic Acids Res.*, **20**, 1117–1122.
16. Yen, P.H., Li, X.M., Tsai, S.P., Johnson, C., Mohandas, T. and Shapiro, L.J. (1990) Frequent deletions of the human X chromosome distal short

arm result from recombination between low copy repetitive elements. *Cell*, **61**, 603–610.
17. Yen, P.H., Allen, E., Marsh, B., Mohandas, T., Wang, N., Taggart, R.T. and Shapiro, L.J. (1987) Cloning and expression of steroid sulfatase cDNA and the frequent occurrence of deletions in STS deficiency: implications for X-Y interchange. *Cell*, **49**, 443–454.
18. Sinclair, A.H., Berta, P., Palmer, M.S., Hawkins, J.R., Griffiths, B.L., Smith, M.J., Foster, J.W., Frischauf, A.M., Lovell-Badge, R. and Goodfellow, P.N. (1990) A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature*, **346**, 240–244.
19. Ma, K., Inglis, J.D., Sharkey, A., Bickmore, W.A., Hill, R.E., Prosser, E.J., Speed, R.M., Thomson, E.J., Jobling, M., Taylor, K. *et al.* (1993) A Y chromosome gene family with RNA-binding protein homology: candidates for the azoospermia factor AZF controlling human spermatogenesis. *Cell*, **75**, 1287–1295.
20. 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. and Page, D.C. (1996) The DAZ gene cluster on the human Y chromosome arose from an autosomal gene that was transposed, repeatedly amplified and pruned. *Nature Genet.*, **14**, 292–299.
21. Lahn, B.T. and Page, D.C. (1999) Retroposition of autosomal mRNA yielded testis-specific gene family on human Y chromosome. *Nature Genet.*, **21**, 429–433.
22. Mardon, G. and Page, D.C. (1989) The sex-determining region of the mouse Y chromosome encodes a protein with a highly acidic domain and 13 zinc fingers. *Cell*, **56**, 765–770.
23. Mardon, G., Luoh, S.W., Simpson, E.M., Gill, G., Brown, L.G. and Page, D.C. (1990) Mouse Zfx protein is similar to Zfy-2: each contains an acidic activating domain and 13 zinc fingers. *Mol. Cell. Biol.*, **10**, 681–688.
24. Kay, G.F., Ashworth, A., Penny, G.D., Dunlop, M., Swift, S., Brockdorff, N. and Rastan, S. (1991) A candidate spermatogenesis gene on the mouse Y chromosome is homologous to ubiquitin-activating enzyme E1. *Nature*, **354**, 486–489.
25. Handley, P.M., Mueckler, M., Siegel, N.R., Ciechanover, A. and Schwartz, A.L. (1991) Molecular cloning, sequence, and tissue distribution of the human ubiquitin-activating enzyme E1. *Proc. Natl Acad. Sci. USA*, **88**, 258–262.
26. Brown, G.M., Furlong, R.A., Sargent, C.A., Erickson, R.P., Longepied, G., Mitchell, M., Jones, M.H., Hargreave, T.B., Cooke, H.J. and Affara, N.A. (1998) Characterisation 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.*, **7**, 97–107.
27. Mitchell, M.J., Woods, D.R., Tucker, P.K., Opp, J.S. and Bishop, C.E. (1991) Homology of a candidate spermatogenic gene from the mouse Y chromosome to the ubiquitin-activating enzyme E1. *Nature*, **354**, 483–486.
28. Soulard, M., Della Valle, V., Siomi, M.C., Pinol-Roma, S., Codogno, P., Bauvy, C., Bellini, M., Lacroix, J.C., Monod, G., Dreyfuss, G. *et al.* (1993) hnRNP G: sequence and characterization of a glycosylated RNA-binding protein. *Nucleic Acids Res.*, **21**, 4210–4217.
29. Wilkinson, G.S., Presgraves, D.C. and Crymes, L. (1998) Male eye span in stalk-eyed flies indicates genetic quality by meiotic drive suppression. *Nature*, **391**, 276–279.
30. Atlan, A., Mercot, H., Landre, C., Montchamp-Moreau, C. (1997) The Sex-ratio in *Drosophila simulans*: geographical distribution of distortion and resistance. *Evolution*, **51**, 1886–1895.
31. Hurst, L.D. (1996) Further evidence consistent with Stellate's involvement in meiotic drive. *Genetics*, **142**, 641–643.
32. Palumbo, G., Bonaccorsi, S., Robbins, L.G. and Pimpinelli, S. (1994) Genetic analysis of Stellate elements of *Drosophila melanogaster*. *Genetics*, **138**, 1181–1197.
33. Hurst, L.D. (1992) Is Stellate a relict meiotic driver? *Genetics*, **130**, 229–230.
34. James, W.H. (1987) The human sex ratio. Part 1: A review of the literature. *Hum. Biol.*, **59**, 721–752.
35. Li, W.H. (1997) *Molecular Evolution.* Sinauer Associates, Sunderland, MA.
36. Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics*, **52**, 1–8.
37. Hudson, T.J., Stein, L.D., Gerety, S.S., Ma, J., Castle, A.B., Silva, J., Slonim, D.K., Baptista, R., Kruglyak, L., Xu, S.H. *et al.* (1995) An STS-based map of the human genome. *Science*, **270**, 1945–1954.